

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 08:09:06 ; Search time 1113 Seconds
(without alignment)

5129.900 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGDKSLTSSYGIQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_pool/US10066521/runat_13072004_122212_11071/app_query.fasta_1.1543
-DB=N Geneseq_29Jan04 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 -CGN_1_1_819 @runat_13072004_122212_11071 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7074	100.0	4035	6 AAL44356	Aal44356 Human PYR
2	5756	81.4	3489	6 AAD41224	Aad41224 Human EMB
3	5756	81.4	3926	7 ADA45218	Ada45218 Human MAT
4	5741.5	81.2	3900	6 ABK48628	Abk48628 Human MAT
5	5741.5	81.2	3900	7 AAD49018	Aad49018 Human MAT
6	5724.5	80.9	5859	6 AAL47135	Aal47135 Pyrin dom
7	5724.5	80.9	6939	6 AAL47131	Aal47131 Pyrin dom
8	5724.5	80.9	6939	6 AAL47140	Aal47140 Pyrin dom

9	5683.5	80.3	3830	7 ADA45220	Ada45220 Human MAT
10	5120	72.4	3226	6 ABX97181	Abx97181 Human NOV
11	2688	38.0	3447	6 ABK48611	Abk48611 Mouse MAT
12	2688	38.0	3447	7 AAD49001	Aad49001 Mouse MAT
13	2005	28.3	1157	6 ABK48609	Abk48609 Human MAT
14	2005	28.3	1157	7 AAD48999	Aad48999 Human MAT
15	1435	20.3	3368	6 AAL44366	Aal44366 Human PYR
16	1434	20.3	2985	6 ADE36408	Ade36408 Human PAA
17	1335	18.9	1075	6 ABK48610	Abk48610 Human MAT
18	1335	18.9	1075	7 AAD49000	Aad49000 Human MAT
19	1331.5	18.8	3300	6 AAL47129	Aal47129 Pyrin dom
20	1320.5	18.7	3190	6 ADE36416	Ade36416 Human PAA
21	1316.5	18.6	3190	7 ADA53511	Ada53511 Human cod
22	1315.5	18.6	3172	6 ABN99366	Abn99366 Human sec
23	1314.5	18.6	2767	7 ABT16018	Abt16018 NOVX rela
24	1303.5	18.4	3186	6 AAL44363	Aal44363 Human PYR
25	1289.5	18.2	2575	6 AAI70683	Aai70683 Human nuc
26	1255	17.7	3857	4 AAD14323	Aad14323 Human PYR
27	1255	17.7	3857	8 ABX93556	Abx93556 Huma cDNA
28	1255	17.7	3857	8 ACD27909	Acd27909 Human PYR
29	1250.5	17.7	3531	6 AAL47128	Aal47128 Pyrin dom
30	1248.5	17.6	3189	5 AAI67185	Aai67185 Nucleotid
31	1238	17.5	2763	6 ABQ75801	Abq75801 Human MDD
32	1234	17.4	3431	4 AAD14322	Aad14322 Human nuc
33	1234	17.4	3431	8 ABX93555	Abx93555 Huma cDNA
34	1234	17.4	3431	8 ACD27908	Acd27908 Human PYR
35	1219.5	17.2	2691	6 AAL44365	Aal44365 Human PYR
36	1217	17.2	3150	5 AAI67184	Aai67184 Nucleotid
37	1189.5	16.8	3306	9 ADC30316	Adc30316 Human nov
38	1143.5	16.2	3069	6 AAL47134	Aal47134 Pyrin dom
39	1128.5	16.0	3263	6 AAI70684	Aai70684 Human nuc
40	1128.5	16.0	3459	6 ABQ78049	Abq78049 Human CGD
41	1068.5	15.1	3085	5 AAS68111	Aas68111 DNA encod
42	1001	14.2	2868	5 AAS33642	Aas33642 cDNA sequ
43	997.5	14.1	2847	4 AAS01487	Aas01487 Human sec
44	997.5	14.1	2847	7 ABZ73494	Abz73494 Secreted
45	997.5	14.1	2847	7 ADA98038	Ada98038 Human sec

ALIGNMENTS

RESULT 1
AAL44356
ID AAL44356 standard; cDNA; 4035 BP.

XX AAL44356;

XX 31-OCT-2002 (first entry)

XX Human PYRIN-5 cDNA sequence.

XX Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-5.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 1..4035
FT /*tag= a
FT /product= "Human PYRIN-5"

XX WO200261049-A2.

XX 08-AUG-2002.

XX 31-JAN-2002; 2002WO-US002967.

XX 31-JAN-2001; 2001US-0265231P.

PR 10-SEP-2001; 2001US-0318645P.
XX (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX
XX Bertin J, Wang W, Blatcher M;
XX P-PSDB; AAO15585.
XX
XX WPI; 2002-627477/67.
XX P-PSDB; AAO15585.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
PT diagnosing stress-related, apoptotic and inflammatory responses, or for
PT treating inflammatory and immune system disorders, cancers, or
PT neurological diseases.
XX
XX Claim 4; Fig 3; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
CC useful for modulating and diagnosing stress-related, apoptotic and
CC inflammatory responses. The PYRIN protein and DNA sequences are useful
CC for treating: inflammatory disorders and immune system disorders (e.g.
CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
CC (e.g. systemic lupus erythematosus and arthritis); and neurological
CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
CC protein and DNA sequences may also be used in screening assays, detection
CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
CC predictive medicine (e.g. diagnostic assays, clinical trials and
CC pharmacogenomics) and transcription profiling. The present DNA sequence
CC encodes the human PYRIN-5 protein
XX

SQ Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4035
Score: 7074.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAL44356 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGAGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCATGCTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40
Db 61 CTGACAAAGGAAGAAATTTACAGCAATTCAGGAATTTACTAAGAGAAGAAATCTTCAGATCG 120
Qy 41 ThrThrCysSerLysProGlnPheGluLeuLeuAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATCGAGATCGCAACGCTGGAAATGCTTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerLysLeuPheGlu 80
Db 181 CTCCTTGATGATGATTATGAGCATCGCTGGCCCTGGGCTACGTCATAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysLysSerGln 100
Db 241 AACATGAACCTGCGAACCCCTCTCGAGAGAGGACGCGGATGACATGAAAAAATTTTCAA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluLeuSerGlnAla 120
Db 301 GCTATGNAACAAGAGGTGCCACAGCAGCAGAGACAGACAGACAAAGAAATTTACAAGCT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThr 140
Db 361 ATGGAAACAAGAGGTGCCACAGCAGCAGCAGACAGACAGAAACAAGACATGAGGTGACACA 420

Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGGACTACAGAGTACGCTGATGACAAATTCGCTGAGGAGGAGGATGTACGTCTGTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAAACACTGCTGCTGACTGGCCGGAATGCAAAACGTTGGCTGGTCTTTTATTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyLysLys 200
Db 541 GACCGGTGGGGCTTCCGGCTCCGACGCTGGTCTTCACGGAAGTCAGGAATTTGGAAA 600
Qy 201 SerAlaLeuAlaArgArgLysValLeuValLeuValLeuValLeuValLeuValLeu 220
Db 601 TCGGCTCTAGCCAGAAAGATCGTGTCTGTGGCGCAAGGTGGACTCTACCAAGGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTTCTCTAGCTCTTCT 720
Qy 241 ThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGluLeuMetSer 260
Db 721 ACAGAGTTTATCTTCCAGGAGTGGCCAGACTTCCAGGCTCCGGTGCAGGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuLeuPheLeuLeuAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAAGAAAGGCTGTGTTCATCATGTACGGTTTCATGACCTGGGCTCTGCTCTCAAC 840
Qy 281 AsnAspThrLysLeuLysCysLysAspTrpAlaGluLysGlnProPheThrLeuLeuArg 300
Db 841 AATGACACAAAGCTCTGCAAAAGATCGGGCTGAGAAGCAGCTCCGTTTCACTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuLeuValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCT 960
Qy 321 GlyThrGluLysLysLysSerGluValValSerProArgTyrLeuLeuValArgGlyLe 340
Db 961 GGCACAGAGAGCTCAAGTCAGAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Qy 341 SerGlyGluGlnArgLysLeuLeuLeuGluArgGlyLysGlyLysGlnLysThr 360
Db 1021 TCCGGGGAACAAAGAAATCCACTTGTCTCTGAGCGGGATGTTGAGCATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaLeuMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGTTGCTGCTGATCATGAACACCGTGAAGTCTCGACAGTGCACAGTGCACCGCC 1140
Qy 381 ValGlySerLeuLeuLysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCTGGCTGAGCTGACAGACGTTGGTGGGGAGAGCGTCCGC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACACCTCACAGGCTGACCGCTCTTGTGTCTTTCATCAGCTCACCCCT 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGGGTGGTCTGGGGCTGTCTCAATCTCGAGGAAGAGTGTCTCTGAAGCGCTTCTGC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet 460
Db 1321 CGTATGGCTGTGAGGAGTGTGGAATAGGAAGTCAAGTGTGTGATGTGACGACCTCATG 1380
Qy 461 ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnLeuLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTCTGTTTTCACATGAACATCTCTTCCCA 1440
Qy 481 AspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACCTTCTTCCACTCTCAGTCTCCAGGACTTCTGTGCC 1500
Qy 501 AlaLeuTyrTyrValLeuGluGlyLeuGluLeuGluProAlaLeuCysProLeuTyrVal 520

Db 1501 GCCTTGACTACGTGTTAGAGGGCTCGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAAGACAAAGAGTCAATGGAGCTTAACAGGAGGCTTCCATATCCACTGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu 560
Db 1621 ATGAAGCGTTCTTGTGTTGGCTCGTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCGTTCCTCGGGGTGAACAGAGCTTTCGCACTGGGTCTCTCTGTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGGAGCTAATGACACCCAGGAGACACCTTGAGCGCTTCCACTGTCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTGTTGCTTGGCAATTAACAGCTTCCAAAGAGTGTGCTT 1860
Qy 621 ProLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 640
Db 1861 CCGATTAAACAGAACCTGGACTTGATAGCATCTTCTCTCTGCTTCCAGCACCTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyLysPheProArgAspGluSerAlaGluAla 660
Db 1921 TTGGGAAATTCGGGTGGAGTCAAGGGAGTCTTCCAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGTCCTCTATGATGCGGAGTAAGACCTCATTTGAGGAGCAGTGGAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTCTGCTCCATGCTTGGCACCCACCCACCTGCGGAGCTGGACCTGGGAGATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 2101 CTGACAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCAATCCACCTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCTGATGTTTAGAATGCACAGATTACCCCTGTTGTCAGCACCTCTGGAGATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCCACCCACCTGAAGGAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuLeuLeuLeuArg 780
Db 2281 GATGTAGGATGGCGTGTGAAGCTTAAACACCCAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGGATTGCTGTGATTGACCCATGCTGTACCTTGAGATCTCCCAATCTTTACGCC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGAAACAAGGTGACAGACCAGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTAGTGTGCTTTGAGAGTCTCCAGTGGCCCTCGCAGAAAGCTGATCTGGAGGAC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCACGGGTTCAGAGTCTGAGCTTCAGCCCTCAGCAACCCGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880

Db 2581 TTGACACACCTGTGCTTATCCAAACAGACCTGGGAAACGAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGTGTAATCAGTGCACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGCTGGCTGTGCTTCTTGCACCTTGCCTTATGGGTAACTCATGGCTGACGCAC 2760
Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuLeuLeuValMet 940
Db 2761 CTGAGCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTCCGAGGTCTG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuGluValLysCysHisLeuThrAlaAla 960
Db 2821 AGAGAACCATCTTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCAGCGCGC 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAATGCCCTGGGTGACGCTGGGGTTGCTGCGCTGTGCGAGGAGCTGAAGCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCACCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCAGTCTTAAACCTGGTGCAAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTGCGCCTTTCGCTGTCACCGTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGlu 1080
Db 3181 TTACAGATTAATGGCTGTGGAAATGGCAGTACCTGTGCAATTAAGGAAGCTGCTGGAG 3240
Qy 1081 GluValGlnLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAAGTCGAGCTACTCAAGCCCGAGTCGTAATGACGCTAGTTGGCATCTTTTGTATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCCGCTCCCTGAAAGCCGGGATGGCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysLysArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTCTGTGGGGATGAACCCAGAGCAGAGAGCGTGTGCTTCTGCTGGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyLysSer 1160
Db 3421 TTCAAGAGCAGTACAGATTTTGGCAAGTCTCTGCTGGCCACCGCAAAATGGTAGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGATTGACAAACGTGGAGCAGAGCTCCCCCAACCCATGGCAGCAGCAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTrpSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTGATGTTGGATATTCGGAGCCTGGTCTGAAACTGCTAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGGCTTGGATCCAAACAGTGTGATCATGACCAAGAGGTATGCGCTGCTCACTAGGGAGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTCTGTATGATGACCAAGCGGTGTCTCTGT 3720

QY 1241 HisTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
 DB 3721 CACTGGAGCGCTGGGCTCTAGGGCTGGTGCTTAACAGTGTCTGATGACCAAGCGGT 3780
 QY 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
 DB 3781 GTGTCCTGGTCACTGGAGCGGCTGGGCTCGAGGGCTTGTGTCCACAGTGTGATGAC 3840
 QY 1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
 DB 3841 CACAGCGTGTGGCTGTCTCACTGGAGCGGCTGGGCTCGAGGGCTTGTGTCCACAGT 3900
 QY 1301 AlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyVal 1320
 DB 3901 GCTGATGACACACAGCGTGTGTCTCTCACTGGAGCGGCTGGGCTCGAGGGCTTGTG 3960
 QY 1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
 DB 3961 TCCACAGTGTGATGACACAGCGGTGTGTCTCTCACTGGAGCGGCTGGGCTCGAG 4020
 QY 1341 GlyLeuValSer 1344
 DB 4021 GGGCTGGTGTCT 4032

RESULT 2

AD41224
 ID AD41224 standard; cDNA; 3489 BP.

AC AD41224;

DT 30-OCT-2002 (first entry)

DE Human EMBRY-1 cDNA.

Human; embryogenesis associated protein; AIDS; reproductive disorder; infertility; endometriosis; endometrial tumour; inflammatory disorder; autoimmune disorder; acquired immune deficiency syndrome; transgenic; ovarian tumour; contact dermatitis; placenta disorder; preeclampsia; EMBRY-1; allergy; gene therapy; gene; ss.

OS Homo sapiens.

PH Key Location/Qualifiers
 CDS 1..3489

FT /*tag= a
 FT /product= "EMBRY-1 protein"

XX WO200248362-A2.

XX 20-JUN-2002.

XX 14-NOV-2001; 2001WO-US043956.

XX 15-NOV-2000; 2000US-0249407P.

XX (INCY-) INCYTE GENOMICS INC.

XX Rankumar J, Arvizu C;

XX WPI; 2002-537629/57.

XX P-PSDB; AAE25053.

PT New polypeptides of human embryogenesis associated proteins for screening
 PT modulators useful for treating or preventing disorders e.g.

PT endometriosis, infertility, allergy, preeclampsia.

PS Claim 58; Page 95-96; 97pp; English.

XX The invention relates to human embryogenesis associated proteins (EMBRY)
 CC and nucleic acid molecules encoding such proteins. EMBRY sequences are
 CC useful for screening modulators useful for treating or preventing
 CC disorders associated with abnormal expression of EMBRY. The disorders

CC treated include reproductive disorders such as infertility,
 CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
 CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,
 CC contact dermatitis; disorders of the placenta such as preeclampsia,
 CC abruptio placentae etc. Sequences of the invention are also useful for
 CC analysing a proteome of a tissue or a cell type. EMBRY proteins are
 CC useful as immunogens for preparing antibodies. Polynucleotides of the
 CC invention are useful for creating knockin humanised animals or transgenic
 CC animals to model human diseases. They are also used in gene therapy. The
 CC present sequence is human EMBRY-1 cDNA

XX
 SQ Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3489
 Score: 5756.00 Matches: 1102
 Percent Similarity: 95.34% Conservative: 2
 Best Local Similarity: 95.16% Mismatches: 0
 Query Match: 81.37% Indels: 54
 DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AD41224 (1-3489)

QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTrpGlu 20
 DB 1 ATGGAAGAGACAAATCGCTCACCTTTCCAGCTACGGCTGCAATGGTGTCTCTATGAG 60
 QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
 DB 61 CTAGACAAGGAAGATTTCCAGACATTCAGAGATTAATAAGAAAGAAATCTTCAGAATCG 120
 QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
 DB 121 ACCACATGCTCTATTCCACAGTTTGAATCGAAGATGCCAAGTGGAAATGTCTGGCAGTC 180
 QY 61 LeuLeuHisGluTrpTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
 DB 181 CTCTTCATGAGTATTATGAGCATCGCTGGCTGGCTACGTCACATAGCATCTTTTGA 240
 QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLys----- 96
 DB 241 AACATGAACCTGCGAACCTCTCGGAGAAGGACGCGGATGACATGAAAAAATTCACAGAA 300
 QY 96 ----- 96
 DB 301 GATCCTGAAGCAACGATGATGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
 QY 96 ----- 96
 DB 361 TATGGCATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAACAC 420
 QY 97 ----- LysIleSerGlnAlaMetGluGlnGluGly 106
 DB 421 AAGTATGTTGGAATTCATTCTCTTTTTCAGAAATTTTCAAGCTATGGAACAAAGAGGT 480
 QY 107 AlaThrAlaAlaGluThrGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAla 126
 DB 481 GCCACAGCAGCAGACAGAGAACAAGAAATTTCAAGCTATGGAACAAGAGTGCC 540
 QY 127 ThrAlaAlaGluThrGluGlnGlnGlyHisGlyVasptThrTrpAspTyrLysSerHis 146
 DB 541 ACAGCAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 600
 QY 147 ValMetThrLysPheAlaGluGluAspValArgArgSerPheGluAsnThrAlaAla 166
 DB 601 GTGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTCGAAAAACACTGCTG 660
 QY 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
 DB 661 GACTGCGCGGAATGCAACGTTGGCTGTGCTTTTGTATTCAGACCGGTGGGCTTCCGG 720
 QY 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206

Db 721 CCTCCACGGTGGTTCTGACGGAAGTCAGAAATTGGGAATCGGCTCTACCCAGAAGG 780
Qy 207 IleValLeuCystrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTGCTGTGCTGGGCGCAAGGTGGAATCTTACCGGGAATGTTCTCTACGTCTTCTTC 840
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCCGGTAGAGAGATGACAGCGGAAGAAGAGAGACAGTGTACAGAGTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCAGGCTCCGGTGACGAGATCATGTCCCCAGCAGAAAGGCTGTG 960
Qy 267 PheIleIleAspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTCATCATGTACGGTTTCGATGACTGGGCTCTGTCTCAACATGACACAAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProProPheThrIleIleArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTGGAAGACAGCTCCGTTCACTCCCTCATACGACGTCTGCTGAGGAAGTC 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTGCTCCCTGAGTCTCTCTGATGTCACCGTCAGAGACGTGGGCGACAGAGAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGlnArgIle 346
Db 1141 TCAGAGGTCTGTCTCCCGTTACTGTGTAGTAGAGGAATCTTCGGGGAAACAAAGATC 1200
Qy 347 HisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGCTCTTGAGCGCGGATTTGGTGAGCATCAGAAAGACACAAAGGGTTCGTCGCATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCys 386
Db 1261 ATGAACAACCGTGAGCTGCTCGACAGTGCAGAGTGGCCGCGCTGCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCCCTGAGCTGCAGGAGTGGTGGGGAGAGCGTTCGCCCTTCACCAAAACGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCTGCACGCGCTTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGTCCGCGCG 1440
Qy 427 CysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCTCAATCTGGAGGAAGAGTTGTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGlu 466
Db 1501 GTGTGGATAGGAAGTCAGTGTGTGACGGTGACGACCTCATGTTCAAGGACTCGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTGTTTCACTGAACATCTCTCTCCAGACAGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeu 506
Db 1621 TACTACACTTTCTCCACTTCAGTCTCCAGAGTTCGTGCGCGCTTGTACTAGCTGTTA 1680
Qy 507 GluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGCGCTGGAAATCGACCGAGCTCTCTGCCCTCTGTACGTTTGAGAGACAAAGAGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPhe 546
Db 1741 ATGGAGCTTAAACAGCGAGGCTTCCATATCCACTTCGCTTTGGATGAAGCGTTTCTGTGTT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGAGCGAAGACGTAAAGAGGCGCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCC 1860

Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAGACTTTCGACTGGGTCTCTCTGTGGGTGAGCAGCTTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606
Db 1921 ACCACCCAGGAGACACCTTGGACGCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeu 626
Db 1981 TTTGTTTCGCTTGGCATTTAAACAGCTTCCAAGAAGTGTGGCTTCCGATTAACCAAGACTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATGACATCTCTCTTCGCTCCAGCACTGTCCGATTTTTCGCGAATAATTCGGGTG 2100
Qy 647 AspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCTCTGTGTCCTCTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATTAAGACCTCATTTAGGAGCAGGTGGGAAGATTTCTGCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCACCCACACCTGCGGACGTGGACCTGGGCGAGCAGCATCTCTGACAGAGCGGCCCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTGTGTGCAAGCTGAGGCATCCCACTGCAAGATACAGACCTGTATTTTAGA 2340
Qy 727 AsnAlaGlnIleThrProGlyValGlnHisIleuTrpArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTACCTCGTGTGTCAGACCTCTGGAGAAATCGTCATGGGCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCys 766
Db 2401 CTAAAGTCCCTCAACTTGGGAGGACCCACCTGAGGAAGAGGATGTAAAGATGGCGTGT 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeu 786
Db 2461 GAAGCTTAAACACCCAAAATGTTTGTGGAGTCTTTGAGGCTGGATTCGTGGGATG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleuThrThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTTACCTGAAGATCTCCAAATCCTTACGACCTCCCGCAGCCTGAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGAGCTTGGCAGGAAACAGGTGACAGACCGGAGTAATGCTCTCAGTGATGCTTGTG 2640
Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTCGCGCTTCGCAAGAGCTGATACTTGGAGGACTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GGTGTCCAGAGTCTGGGCTCAGGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCTTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCAAACACAGCTTGGGGAACGAGGTGTAAATCTACTGTGTGATCCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGAGAGGCTGATGCTGAATCAGTGCACCTGAGACCGGCTGGCTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGCACCTTTCGCTTATGGGTAATCTCATGTGTCGACCGACCTGAGCCTTAGCATGAAC 2940

```
QY 927 ProValGluAspAenGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
DB 2941 CCTGTGGAGACATGGGTGAGCTTCTGTCCGAGGTCTCATGAGAACCATCTTGTCAT 3000
QY 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
DB 3001 CTCAGGACCTGGAGTGGTAAAGTGTCTATCCACCGCGCGTCTGTGAGAGTCTGTCC 3060
QY 967 CysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
DB 3061 TGTGTGATCTCGAGAGCAGACACCTGAAGAGCCTGGATCTCACGGACAATGGCCCTGGT 3120
QY 987 AspGlyGlyValAlaAlaLeuCysGluLysGlnLysAsnSerValLeuThrArg 1006
DB 3121 GACGGTGGGGTGTGCACTGTGGAGGACTGAGCAAAAGAACAGTGTTCGACGAGA 3180
QY 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
DB 3181 CTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGTGAGGCACCTCTCCCTGGCCCTT 3240
QY 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
DB 3241 TCCTGCAACCGGCATCTGACAGCTCTAAACCTGTGTGCAGATAACTTCAGTCCCAAGGA 3300
QY 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleGlyLeu 1066
DB 3301 ATGATGAAGCTGTGTTCGGCTTTTGCCTGCCACGTCTAACTTACAGATAATTTGGGCTG 3360
QY 1067 TrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGluValGlnLeuLys 1086
DB 3361 TGGAAATGGCAGTACCTCTGCAAAATGAAGAGCTGTGTGGAGGAAGTGCAGTACTCAAG 3420
QY 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspArgHis 1104
DB 3421 CCCCAGTCGTAATGACGGTAGTTGGCAATCTTTTGTATGAAGATGACCGGTAC 3474
RESULT 3
ADA45218
ID ADA45218 standard; DNA; 3926 BP.
XX
AC ADA45218;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human MATER DNA.
XX
KW human ; MATER; maternal antigen that embryos require; NTP-ase;
KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
KW ovarian dysfunction; autoimmune premature ovarian failure;
KW autoimmune disease; contraceptive; autoantibody; ds; gene.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /tag= a
FT /product= "MATER"
XX
FN EP1285964-A2.
XX
PD 26-FEB-2003.
XX
PF 12-JUL-2002; 2002EP-00090246.
XX
PR 10-AUG-2001; 2001DE-01039874.
XX
PA (SCHD ) SCHERING AG.
XX
PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
XX
DR WPI; 2003-302814/30.
DR P-PSDB; ADA45219.
```

```
XX
PT New nucleic acid for human maternal antigen that embryos require protein,
PT useful e.g. in diagnosis and treatment of female infertility.
XX
PS Claim 1; Page 9-11; 31pp; German.
XX
CC This invention describes a novel human MATER (maternal antigen that
CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
CC associated with apoptosis, a defect in MATER activity causes growth
CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
CC MATER gene is found predominantly in ovary, testis and placenta. The
CC products of the invention have antiinfertility, contraceptive,
CC antiinflammatory, immunosuppressive and gynaecological activity.
CC Effectors of the MATER polypeptide are used to treat infertility
CC associated with endometriosis and also ovarian dysfunction, autoimmune
CC premature ovarian failure, inflammation, autoimmune diseases and female
CC infertility, and as contraceptives. Measurements of autoantibodies
CC against MATER in body samples or MATER protein or mRNA in eggs is useful
CC for diagnosis of female infertility, particularly by detecting mutations
CC in the MATER gene with a DNA chip. This sequence encodes the human MATER
CC polypeptide described in the invention.
XX
SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 5756.00 Matches: 1102
Percent Similarity: 95.34% Conservative: 2
Best Local Similarity: 95.16% Mismatches: 0
Query Match: 81.37% Indels: 54
DB: 7 Gaps: 1

US-10-066-521-6 (1-1344) x ADA45218 (1-3926)
QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTrpGlyLeuGlnTrpCysLeuTrpGlu 20
DB 1 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGCTGCAATGGTGTCTCTATGAG 60
QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysLysSerSerGluSer 40
DB 61 CTAGACAAGGAAGAAATTCAGACATTCAGAGGAATTACTAAGAGAAATCTTCAGAAATCG 120
QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
DB 121 ACCATATGCTCTATTCCAGATTTGAAATCGAAGATGCCACGTGGAAATGCTCGGCATC 180
QY 61 LeuLeuHisGluTrpTrpGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
DB 181 CTCTTGCATGAGTATTATGAGCATCGCTGGCTGGCTAGCTCCATTAGCATCTTTTGA 240
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLys----- 96
DB 241 AACATGAACCTGCGAAACCCCTCTCGGAGAGGCGCGGATGACATGAAAAAATTCACGAG 300
QY 96 ----- 96
DB 301 GATCCTGAACACAGATGACTGACCAAGGACCAAGCAAGAAAAAGTCCAGAAAAATAA 360
QY 96 ----- 96
DB 361 TATGGCATGACTAAGCTTATCTTTGGGGTGTCTGACATCTCTGACTCGAATAATAAACAC 420
QY 97 -----LysIleSerGlnAlaMetGluGlnGly 106
DB 421 AAGTATGTTGGAATTCATTCTTTTTCAGAAATTTTCAAGCTATGGAACAAGAGGT 480
QY 107 AlaThrAlaAlaGluThrGluGlnGluIleSerGlnAlaMetGluGlnGlyAla 126
DB 481 GCCACAGCAGCAGACAGACAGAAAGAAATTTCAAGCTATGGAACAAGAGGTGCC 540
QY 127 ThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTrpLysSerHis 146
DB 541 ACAGCAGCAGACAGAGAAGAAACAAAGGACATGGAGGTGACACATGGGACTACAAGAGTCC 600
```



```
Db 2761 TCCACACAGCCTGGGAAACGAGGTGTAATCTACTGTGTCCATCCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGCTGATGCTGAATCAGTGCCACCTGGACACGGCTGGCTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGACCTTGGCTTATGGGTAACCTCATGGCTGACGCACCTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGGAAGCAATGGCGTGAGCTTCTGTGCGAGGTCAATGAGAACCACTGTGTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
Db 3001 CTCCAGGACCTGGAGTTGGTAAGTGTCTATCTACCGCGCGTCTGTGAGAGTCTGTCC 3060
Qy 967 CysValLysSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGAGCAGACACTGAAGAGCCTGGATCTCACGGACAATGGCCCTGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GACGGTGGGGTGTGCTGCACTGTGCGAGGACTGAAGCAAGAACAGATGTTCTGACGAGA 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGTGAGGCACCTCTCTGGCCCT 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnAspSerProLysGly 1046
Db 3241 TCCTGCAACCGGCATCTGACAGTCTAAACCTGTGTGAGAGATACTTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnLysGlyLeu 1066
Db 3301 ATGATGAAGCTGTGTGGCCCTTGGCTGTCCAGCTTAACCTTACAGATAATTGGGCTG 3360
Qy 1067 TrpLysTrpGlnTrpProValGlnLeuArgLysLeuLeuGluValGlnLeuLys 1086
Db 3361 TGGAAATGGCATGACCTGTGCAAAATAGGAAGCTGTGGAGGAAGTGCAGTACTCAAG 3420
Qy 1087 ProArgValValLysAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3421 CCCCAGTCGTAATTGACGGTAGTGTGGCATCTCTTTGATGAAGATGACCCGGTAC 3474
```

RESULT 4

```
ABK48628 standard; cDNA; 3900 BP.
XX ID ABK48628;
AC ABK48628;
XX DT 13-AUG-2002 (first entry)
XX DE Human MATER cDNA.
XX KW Human; gene; ss; contraceptive; antinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX OS Homo sapiens.
XX FH Key
XX CDS 1..3603
XX FT Location/Qualifiers
XX FT /*tag= a
XX FT /product= "Human MATER"
XX FN WO200232955-A1.
XX PD 25-APR-2002.
XX
```

PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
XX WPI; 2002-454595/48.
DR P-PSDB; AAU79526.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 11; Page 80-85; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA.
CC Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
CC published human genomic DNA sequences, while the remainder of the
CC sequence was determined by direct cloning of human ovarian cDNAs
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x ABK48628 (1-3900)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 154 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTCTCTATGAG 213
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysLysSerSerGluSer 40
Db 214 CTAGAAGAAGAAATTTACAGATTTCAAGGAATTACTAAGAAGAAATCTTCAGATCG 273
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTCCACAGTTTGAATCGAGAATGCCAATGCGAATGTCTGGCAGTC 333
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 334 CTCTTCATGAGTATTATGAGCATCGCTGGCCCTGGCTACGTCCTCATTAGCATCTTTGAA 393
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
Db 394 AACATGAACCTGCAACCCCTCTCGGAGAGGCAAGGATGACATGAAGACATTCACCA 453
Qy 97 ----- 97
Db 454 GAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCCAGGAATT 513

QY 98 -----IleSer 99
Db 514 TCACAAGCTGTGCAACAAGATAGTCCACAGCTGCAGAGACAAAGAAATTTCA 573
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnIleSerGln 119
Db 574 CAAGCTATGGAAACAAGAAGTCCACAGCAGCAGAGACAGAAGAAACAAGAAATTTCAAA 633
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnIleHisGlyAsp 139
Db 634 GCTATGGAAACAAGAAGTCCACAGCAGCAGACAGACAGAAACAGAGCATGGAGTGAC 693
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
Db 694 ACATGGGACTACAAGAGTCACGTGATGACCAAAATTCGCTGAGGAGGAGATGACGTGCT 753
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 754 AGTTTTGAAACACATGCTGCTGACTGGCCGGGAAATGCAAAACGTGGCTGGTCTTTGAT 813
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCGGGCTTCGACGGTGGTTCCTGCACGGAAAGTCAGGAATTGGG 873
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly 219
Db 874 AAATCGGCTCTAGCCAGAAAGATCGTGTCTGGCGCGAAGGTGGACTCTTACCAGGGA 933
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACGTCTTCTTCCTCCCGTTAGAGAGATGCAGCGGAAGAGAGACAGT 993
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCACAGAGTTTCATCTCAGGGAGTGGCCAGACTCCAGGGCTCCGGTACCGAGATCATG 1053
QY 260 SerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu 279
Db 1054 TCCCGACAGAAAGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGTCTCT 1113
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
Db 1114 AACATATGACACAAAGCTCTGCAAGACTGGGCTGAGAGAGCCCTCCGTTACCCCTCAT 1173
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGTCTGCTGAGAAAGTCTGTCTCTGATGCTCTTCTGATGCTCACCGTCAGAGAC 1233
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTCTCTCCCGTTACCTCTGTTAGAGGA 1293
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGACAAAGAAATCCACTTGTCTCTTGTAGCGCGGGATTGTTGAGCATCAGA 1353
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACACAAGGGTGTGTCGCATCATCAACAACCGTGGCTGCTGACCCAGTCAGGTGCC 1413
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
Db 1414 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAGAGCGTC 1473
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCCCCCTTCAACCAACCGCTCACAGGCTGCACGCCGCTTTTGGCTTTTCATCAGCTCACC 1533
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLysArgPhe 439
Db 1534 CCTCAGGCGGTGTCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTCCTGAAGCGCTTC 1593

QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
Db 1594 TGCCTATGGCTGTGGAGGGAGTGTGGAATAGGAAGTCAGTGTGTGATGGTGACACCTC 1653
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGGTTCAAAGACTCGGGGAGTCTGAGCTCGTCTCTGTTTCACATGAATCATCTCTC 1713
QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCAGACAGCCACTGTGAGGAGTACTACCTTCTTCCACCTCAGTCTCCAGAGACTTCTGT 1773
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
Db 1774 GCCGCTTGTACTACGTGTAGAGGGCTGGAATCGAGCCAGCTCTGTGCCCTCTGTAC 1833
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAAAGACAAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCCATATCCACTCGCTT 1893
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAAGAGGCCACTGGAGGTC 1953
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCCCGTTCCCTTGGGGTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTG 2013
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTGAGCAGCGCTTAATGCCACCACCCAGGAGACACCTTGGACGGCTTCCACTGTCTT 2073
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACGCTTCCAAAGAGTGTGG 2133
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAAACAGAACCTGGACTTGATAGCATCTTCTTCTGCTCCAGCACCTGTCCG 2193
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTTCGGGAAAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGATGCCGCTGAG 2253
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
Db 2254 GCATGCTCTGTGTCCTCTATGGATCGGGATTAAGACCTCATTTAGGAGCAGTGGGAA 2313
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCTGCTCCATGCTTGGCACCCACACCTCGCGCAGCTGGACCTGGGCAGCAGC 2373
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGAGCATCCACCTGCAAG 2433
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGTGATTTAGAAATGACAGATTACCCCTGGTGTGCAACACCTCTGAGGA 2493
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2494 ATCGTCATGGCCAAACCGTAACCTAAGATCCTCAACTTGGGAGGCGCCACCTGGAAGGA 2553
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGAAGGATGGCGTGTGAAGCCTTAAACACACCCAAAATGTTTGTGGAGTCTTTG 2613
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGATTTGCTGTGGATTGACCCGCTGTACCTGAAGATCTCCCAAAATCTTTACG 2673
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819

```
Db 2674 ACCTCCCCAGCCCTGAATCTCTGAGCTGCGAGAAACAGGTGCAGACCCAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2734 ACGCTCTCAGTGATGCTTCCAGGGTCTCCAGTGGCCCTGCAGAAAGCTGATACAGG 2793
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCACAGCCAGCGGTGGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACCTGTGGCTATCCACACAGCCCTGGGGAACGAGGTGTAAATCTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCAC 2973
Qy 900 LeuAspThrAlaGlyCysGlyPheIleuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGGACAGCGTGCTGGTGTCTTTCACCTTCGCTTATGGGTAACTCATGGCTGACG 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 3034 CACCTGAGCCTTAGCATGAACCTGTGGAGACAAATGGCGTGAAGCTTCTGTGGCAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 3094 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTCGAGTTGGTAAAGTGTCTATCTCACCGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 CGCTGCTGTGAGAGTCTGCTCTGTGATCTCGAGGAGCAGACACCTGAAAGAGCCTGGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValalalalaLeuCysGlyGlyLeuLysGln 999
Db 3214 CTCACGGCAATGCCCTGGGTGACGGTGGGTGCTGGCTGTCCGAGGAGCTGAACAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGGACTGACTTCTGATTGCTGT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 3334 GAGGCACCTCTCTTGGCCCTTCTCTGCAACCCGCACTTGACCACTCTAAACCTGGTGCG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 ATTAACCTCAGTCCCAAGGAATGATGAGCTGTGTCGGCTTGTCCACGCTCT 3453
Qy 1060 AsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATAATTGGCTGTGAAATGGCAGTACCCCTGTGCAATAAAGGAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTGACGGTAGTGTGGCATCTTTTGTAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
```

RESULT 5

AAD49018
ID AAD49018 standard; cDNA; 3900 BP.

XX
XX AAD49018;

AC
DT 07-MAR-2003 (first entry)

XX
XX Human MATER cDNA.

Human; MATER protein; infertility; fertility; contraceptive agent;
gene therapy; gene; ss.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX CDS 1..3603
FT /*tag= a
FT /product= "Human MATER protein"
XX
XX WO200281492-A1.
XX
XX 17-OCT-2002.
XX
XX 29-MAR-2002; 2002WO-US009776.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2003-058494/05.
XX
XX P-PSDB; AAE31749.
XX
XX New isolated variant MATER proteins and nucleic acids, useful for
XX diagnosing, prognosing and treating infertility and reduced fertility,
XX and as contraceptive agents.
XX
XX Example 2; Page 93-98; 110pp; English.
XX
XX The present invention relates to novel MATER proteins and polynucleotides
XX encoding such proteins. The MATER proteins are essential to fertility.
XX Sequences of the invention are useful for diagnosing, prognosing and
XX treating infertility, reduced fertility and as contraceptive agents. They
XX are also useful in gene therapy. The method is useful for detecting a
XX predisposition to or pre-symptomatic screening of an individual for
XX infertility or reduced fertility. The present sequence is human MATER
XX cDNA

SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3900
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: Gaps: 1

US-10-066-521-6 (1-1344) x AAD49018 (1-3900)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTrpGlu 20
Db 154 ATGGAAGGACAAATCGCTCACCTTTCAGCTACGGGCTGCAATGGTCTCTATGAG 213
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40
Db 214 CTAGACAAGGAAGAAATTCAGACATTCAGGAATTAAGGAAGAAATCTTCAGAAATCG 273
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTTCACAGTTTGAATCGAATGCGAATGCCAACGTGGAATGCTTGGCACTC 333
Qy 61 LeuLeuHisGluTrpTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 334 CTCTTCATGAGTATTATGAGGACATCGCTGGCTGGCTACGTCATTCATCTTTGAA 393
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
Db 394 AACATGAACCTGCGAAACCTCTCTCGGAGAGGACCGGGATGACATGAAGACATTACCA 453
Qy 97 ----- 97

QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
 DB 2674 ACCCTCCCCAGCCTGAATCTCTGAGCCTGCAGGAAACAGAGGTGACAGACCAGGAGTA 2733
 QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
 DB 2734 ACGCCTCTCACTGATGCTTCTGAGGCTCTCCAGTGGCCCTGCAGAGCTGATACTGAG 2793
 QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 DB 2794 GACTGTGGCATCACAGCCACGGGTGGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
 QY 860 SerLeuThrHisLeuLysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
 DB 2854 AGCTTGACACACCTGTGCTCTATCAACACAGCCCTGGGGAACGAGGTGTAAATCTACTG 2913
 QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
 DB 2914 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGCTGATGCTGAATCAGTGCAC 2973
 QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
 DB 2974 CTGGACACGGCTGGCTGTGCTTCTTGCACTTGCGCTTATGGGTAACCTCATGGCTGAGC 3033
 QY 920 HisLeuSerLeuSerMetAsnProValGluAspGlnGlyValLysLeuLysLeuVal 939
 DB 3034 CACCTGAGCCCTTAGCATGACCCCTGTGGAACAACTGGCGTGAAGCTTCTGTGCGAGGTC 3093
 QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
 DB 3094 ATGAGAGAACCATCTTGTCACTCCAGGACCTGGAGTTGGTAAAGTGTCACTCACCGCC 3153
 QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
 DB 3154 CGCTGCTGTGAGAGTCTGTCTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 3213
 QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaLeuLysGlyValLysGln 999
 DB 3214 CTCAGGCAATGCCCTGGGTGACGGGTGGGTGCTGCGCTGTCGAGGACTGAAGCAA 3273
 QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
 DB 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGAGCTGACTTCTGATTGCTGT 3333
 QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
 DB 3334 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTTAAACCTGGTGCAG 3393
 QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
 DB 3394 AATACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGGCTTGTGCTGTGCCACGTCT 3453
 QY 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079
 DB 3454 AACTTACAGATAATTGGCTGTGAAATGGCAGTACCTCTGTGCAATPAGGAAGCTGCTG 3513
 QY 1080 GluGluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAsp 1099
 DB 3514 GAGGAAGTGCAGCTACTCAAGCCCGAGTCGTAAATGACGGTAGTGTGGCATCTCTTTTGT 3573
 QY 1100 GluAspAspArgHis 1104
 DB 3574 GAAGATGACCGGTAC 3588

RESULT 6

AAL47135

ID AAL47135 standard; DNA; 5859 BP.

XX AC AAL47135;

XX DT 20-AUG-2002 (first entry)

XX

DE XX Pyrin domain containing protein NALP8/Py12 coding sequence.
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antiporiatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteopathic; neutropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis; gene; ds.

Unidentified.

OS XX

PN W0200240668-A2.

XX XX

PD 23-MAY-2002.

XX XX

PF 30-OCT-2001; 2001WO-EP012545.

XX XX

PR 15-NOV-2000; 2000DE-01056687.

PR 30-NOV-2000; 2000DE-01059595.

XX XX

PA (APOT-) APOTEC RES & DEV LTD.

XX XX

PI Tschoep J, Martinon F;

XX XX

DR WPI; 2002-427093/45.

DR P-PSDB; AAO17863.

XX XX

PT New DNA encoding protein with pyrin domain, useful for treating diseases
 involving impaired signal transduction, particularly inflammation, also
 proteins and antibodies.

XX XX

PS Claim 5; Fig 1; 116pp; German.

XX XX

CC The present invention relates the DNA and their encoded proteins, where
 the proteins contain at least one PYD (pyrin) domain. These can be used
 to treat diseases associated with impaired intracellular signal
 transduction, particularly inflammation such as psoriasis,
 arteriosclerosis, bacterial or viral infections (particularly meningitis
 and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 and Parkinson's diseases. The present sequence is a coding sequence of
 the invention

XX XX

SQ Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 5859

Score: 5724.50 Matches: 1106

Percent Similarity: 90.53% Conservative: 3

Best Local Similarity: 90.29% Mismatches: 7

Query Match: 80.92% Indels: 109

DB: 6 Gaps: 3

US-10-066-521-6 (1-1344) x AAL47135 (1-5859)

QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrglyLeuGlnTrpCysLeuTyrglu 20

DB 1975 ATGGAAGGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGGTGTCTCTATGAG 2034

QY 21 LeuAspLysGluCluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40

DB 2035 CTAGACAAGAGAGATTTCAGACATTCAAGGAATTACTAAGAGAAATCTTCAGAAATCG 2094

QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60

DB 2095 ACCACATGCTCTATTCCACAGTTTGAATCGAGATGCCACGTGGAATGTCTGGCAGCTC 2154

QY 61 LeuLeuHisGluTyrrTyrglyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80

DB 2155 CTCTTCATGAGTATTATGGAGCATCGCTGGGCTACGTCACATAGCATCTTTTGA 2214

QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys----- 97


```
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 4435 AGGCTGGATTCTGTGGATTGACCATGCTGTTACCTGGAAGATCTCCCAATCTTTACG 4494
QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 4495 ACCTCCCCAGCTGMAATCTCTGAGCTGCGAGGAAACAGGTGACAGACCAGGAGTA 4554
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 4555 ATGCTCTCTCAGTGATGCTTGGAGAGTCTCCAGTGGCGCTGCGAGAGCTGATACCTGGAG 4614
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 4615 GACTGTGGCATCACAGCCAGCGGTGGCAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGG 4674
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 4675 AGCTTGACACACCTGTGCTATCCACACAGCCCTGGGGAACGAGGTGTAAATCTACTG 4734
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 4735 TGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCAC 4794
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 4795 CTGACACGGCTGGCTGGTGGTCTTCTGCACTGGCTTATGGGTAACTCATGGCTGACG 4854
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 4855 CACCTGAGCCTTAGCATGAACCTGTGGAGACATGGCGTGAAGCTTCTGTGGAGGTC 4914
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 4915 ATGAGAGAACCATCTTGTCTATCTCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCC 4974
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 4975 GCGTGCTGTGAGAGTGTCTGTGTGATCTGAGAGGAGACACCTGAAGAGGCTGGAT 5034
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 5035 CTCAGGACATGCCCTGGGTGACGGTGGGTGCTGCACTGTGCGAGGACTGAACAA 5094
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 5095 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGT 5154
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 5155 GAGGCACTCTCTTGGGCCCTTCTGCAACCGGCATCTGACCATCTGAACCTGTAACCTGGTCAG 5214
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 5215 AATACTTCATCCCAAGGAATGATGAGTGTGTTGGGCTTGGCTGTGCCACCTCT 5274
QY 1060 AsnLeuGlnIleIleGly----- 1065
Db 5275 AACTTACAGATAATGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAATCTGCTCTA 5334
QY 1065 ----- 1065
Db 5335 GTAGTTTGTCTAAAGTCATCTGTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 5394
QY 1065 ----- 1065
Db 5395 CTGTACCTCCCAACCCCTATAACACGACCGGCACAGAGACTCTGAGGAAGGAATTCAT 5454
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 5455 GGATGGAGTGAAGGCTGTGGAAATGGCAGTACCTGTGTCAAATAAGGAAGCTGCTGGAG 5514
```

```
QY 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 5515 GAAGTGAGCTACTCAAGCCCCGAGTCGTAATGACGGTAGTGTGGCAATCTTTTGATGAA 5574
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 5575 GATGACCGATTGGATCTTCAAAAGCCAGCAACAGTCACTCAGCAAGACAGACTTACAAT 5634
QY 1112 LeuProGluSerArg 1116
Db 5635 CTCATGCGCATCAG 5649
RESULT 7
AAL47131
ID AAL47131 standard; DNA; 6939 BP.
XX
AC AAL47131;
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP5/Py8-hs coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN WO2002040668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01058687.
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTECH RES & DEV LTD.
XX
PI Tschopp J, Martinon F;
XX
DR WPI; 2002-427093/45.
DR P-PSDB; AAO17859.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
```

```
US-10-066-521-6 (1-1344) x AAL47131 (1-6939)
QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTTPCysLeuTyrGlu 20
DB 3055 ATGGAGGAGACAAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 3114
QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
DB 3115 CTAGACAGGAAGAATTTTCAGACATTCAGGAAATTTACTAAAGAAGAATCTTCAGAAATCG 3174
QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
DB 3175 ACCACATGCTCTATTTCCACAGTTTGAATTCGAGAATGCCAACGTGGAAATGTCTGGCCTC 3234
QY 61 LeuLeuHisGluTyrTyrGlyValAserLeuAlaThrSerIleSerIlePheGlu 80
DB 3235 CTCCTTGATGATGATTTATGGAGCATCGCTGGCCCTGGGCTACGTCCATTAGCATCTTTGAA 3294
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
DB 3295 AACATGAACCTCGCAACCTCTCGGAGAGGCGACGGGATGACATGANAAGACATTCACCA 3354
QY 97 ----- 97
DB 3355 GAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGGAAAAAAGTGCCAGGAATT 3414
QY 98 -----IleSer 99
DB 3415 TCACAGCTGTGCACAAAGATAGTGCCACAGCTCGACAGACAAAGAACAAAGAAATTTCA 3474
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln 119
DB 3475 CAAGCTATGGAACAAAGAGGTGCCACAGCAGCAGACAGAGACAGAAAGAAATTTCAAA 3534
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyAsp 139
DB 3535 GCTATGGAACAAAGAGGTGCCACAGCAGCAGACAGACAGAAAGAACAGACATGGAGGTGAC 3594
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
DB 3595 ACATGGGACTACAGAGTACGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGT 3654
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
DB 3655 AGTTTGAACACATGCTGCTGACTGGCCGGAATGCAACGTTGGCTGGTCTTTTGAT 3714
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
DB 3715 TCAGACCGGTGGGGCTTCGGGCTCGCACGGTGTCTGACCGGAAAGTCAGGAATTTGGG 3774
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly 219
DB 3775 AAATCGGGCTTAGCCAGAAAGATCGTGTGCTGGCGCGCAAGGTGGACTCTACCCAGGGA 3834
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
DB 3835 ATGTTCTCTACGCTTCTTCTCCCGGTTAGAGAGATGCGACGCGAAGAGAGAGCAGT 3894
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
DB 3895 GTCACAGATTCATCTCCAGGGAGTGCCAGACTCCCGAGGCTCCGGTGACGGAGATCATG 3954
QY 260 SerArgProGluArgLeuPheIleAspGlyPheAspAspLeuGlySerValLeu 279
DB 3955 TCCGACACAGAAAGCTGTGTTTCATCATTCACGGTTTCGATGACCTGGGCTCTGCTCTC 4014
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
DB 4015 AACATGACACAAAGCTCTGCAAGACTGGGCTGGAAGAGCGCTCGCTCACCCCTCATA 4074
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
DB 4074 GCATGTCTGTGGTCCCTCTATGGATGCGGGATAAGACCCCTCATTTGAGGAGCAGTGGGAA 5214
```

```
DB 4075 CGCAGTCTGCTGAGGAAGGTCTCTCTCCCTGAGTCTCTCTGATCGTCACCGTCAGAGAC 4134
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
DB 4135 GTGGGCACAGAGAGCTCAAGTCAAGTCAAGTCTCTCTCCCGTACCTTGTAGTTAGAGGA 4194
QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 4195 ATCTCGGGGAAACAAAGAAATCCACTTGTCTCTTGGCGCGGATTTGGTGAGCATCAGAA 4254
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
DB 4255 ACAAAAGGGTTGGTGGCATCAACCAACCGTGAGCTGCTCGACCAAGTCCAGGTCGCC 4314
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
DB 4315 GCGTGGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAGAGCGTC 4374
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 4375 GCCCCTTCAACCAACCGCTCACAGGCTGCACGCGCTTTTGTGTTTCATCAGTCACC 4434
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
DB 4435 CCTCGAGGCGTGTTCGGCGCTCTCTCAATCTGGAGGAAAGAGTTGCTCGAAGCGCTTC 4494
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
DB 4495 TGCGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCAGGCTGACGCTTC 4554
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 4555 ATGTTCAAGGACTCGGGGAGTCTGAGTCCGTGCTCTGTTTCACATGAACATCCTTCTC 4614
QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 4615 CCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGT 4674
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 4675 GCGCCTTGTTACTGCTGTGTAGAGGCGCTGGAATTCGAGCCAGCTCTCTGCCCTCTGTAC 4734
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 4735 GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGACAGGCTTCCATATCCACTCGCTT 4794
QY 540 TrpMetLysArgPhePheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 4795 TGGATGAAGCGTTTCTTGTGTGGCCTCGTGAGGAAAGACGTAAGAGGAGCCACTGGAGGTC 4854
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 4855 CTGCTGGGCTGTCCCGTTCCCTTGGGGGTGAAGCAGAAAGCTTCTGCACTGGGTCTCTCTG 4914
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGTTCAGCAGCCTAATGCCCACACCCAGGAGACACCTTGGACGCTTCCACTGTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAAACAGCTTCCAAAGAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCCGATTAACAGAACCTTGGACTTGTATAGCATCTTCTTCTGCTCCAGCAGCTGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTGGGAAATTCGGGTGGATGTCAAGAGGATCTTCCCAAGAGATGATGTCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCTGTGGTCCCTCTATGGATGCGGGATAAGACCCCTCATTTGAGGAGCAGTGGGAA 5214
```

```
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 5215 GATTTCGTCTCATGCTTGGCACCACCACCCACCTCGCGGAGCTGGACCTGGGGCAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLeuThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 5275 ATCTGTACAGACGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAG 5334
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 5335 ATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGA 5394
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 5395 ATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACCTTGGAGGACCCACCTGAAGGAA 5454
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAAGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTTGTGGAGTCTTTG 5514
QY 780 ArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATTGCTGTGGATTGACCATGCTCTTACCTGAAGATCTCCCAAAATCCTTACG 5574
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCCAGCCCTGAATCTCTCAGCGCTGGCAGGAACAGGTGACAGACCCAGGGAGTA 5634
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 5635 ATGGCTCTCAGTGATGCTTGAGAGTCTCCAGTGGCCCTGCGCAGAGCTGATCTGGAG 5694
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGGCATCACAGCCACGGTGTGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCCG 5754
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGlyValAsnLeu 879
Db 5755 AGCTTGACACACCTGTGCTTATCAACACACAGCTGGGGACGAAGGTGTAATCTACTG 5814
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTGATCCATGAGCTTCCCATCTAGTCTGAGAGGCTGATGCTGATCAGTGCAC 5874
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGGACACGGCTGGCTGTGGTCTTCTTGCACTTGGCTTATGGGTAACCTCATGGCTGACG 5934
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuCysGluVal 939
Db 5935 CACCTTGAGCCCTTAGCATGAACCCCTGTGAAGACAATGGCGTGAAGCTTCTGTGGAGGTC 5994
QY 940 MetArgGluProSerCysHisLeuGluAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCC 6054
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 CGGTGCTGTGAGAGTCTGTCTGTGTATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 6114
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCACGACAAATGCCCTGGGTGACGGTGGGTGCTGCTGCTGCTGCGAGGAGCTGAAGCAA 6174
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AAGAACAGTGTCTGACGAGACTCGGGTGGAGCATGTGGACTGACTTCTGATTGCTGT 6234
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGCACTCTCTTGGGCCCTTTCTCTGCAACCGGCATCTGACCACTGTAAACCTGGTGACG 6294
```

```
QY 1040 AsnAsnPheSerProLysGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTTGGCTGTGCCAGTCT 6354
QY 1060 AsnLeuGlnIleIleGly----- 1065
Db 6355 AACTTACAGATATTGGCAATGACTCTGAAGAAAATGACGCTTCTTCGAGAATCTGCTCTA 6414
QY 1065 ----- 1065
Db 6415 GTAGTTTTCCTTAAAGTCACTGTTTCCAGAACCTTATCAATGACATTAAAGGAGAACTTA 6474
QY 1065 ----- 1065
Db 6475 CTGTACTCTCCCAAAACCCCTATTAAACAGACCCCGCACAGAGACTCTGAGGAGGAATTCAT 6534
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GGATGGACTGAAGGGCTGTGGAATGGCAGTACCTCTGCAAAATAAGGAAGCTGTGGAG 6594
QY 1081 GluValGlnLeuLeuLysProArgValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GAAGTGCAGCTACTCAAGCCCGAGTCGTAATTGACGCTAGTTGGCATCTCTTTTGATGAA 6654
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATGGATCTTCAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
QY 1112 LeuProGluSerArg 1116
Db 6715 CTCATGGCATCACAG 6729
RESULT 8
AAL471140
ID AAL471140 standard; DNA; 6939 BP.
XX AC AAL471140;
XX XX
DT 20-AUG-2002 (first entry)
XX DE
XX PYRIN domain containing protein NALP13/Py17 coding sequence.
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; neotropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX OS
XX UNidentified.
XX WO200240668-A2.
XX XX
XX 23-MAY-2002.
XX XX
XX 30-OCT-2001; 2001WO-EP012545.
XX XX
XX 15-NOV-2000; 2000DE-01056687.
XX 30-NOV-2000; 2000DE-01059595.
XX XX
XX (APOT-) APOTECH RES & DEV LTD.
XX TSchopp J, Martinon F;
XX WPI; 2002-427093/45.
XX P-PSDB; AAO17868.
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
XX PS Claim 5; Fig 1; 116pp; German.
```

XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PVD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention

XX Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6939
 Score: 5724.50 Matches: 1106
 Percent Similarity: 90.53% Conservative: 3
 Best Local Similarity: 90.29% Mismatches: 7
 Query Match: 80.92% Indels: 109
 DB: 6 Gaps: 3

US-10-066-521-6 (1-1344) x AAL47140 (1-6939)

Qy	1	MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20
Db	3055	ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG	3114
Qy	21	LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer	40
Db	3115	CTAGCAAGGAAGAATTTTCAGACATTCAGGAATTTACTAAAGAAGAATCTTCAGAAATCG	3174
Qy	41	ThrThrCysSerIleProGlnPheGluIleGluAlaAlaSerValGluCysLeuAlaLeu	60
Db	3175	ACCACATGCTCTATTCCACAGTTTGAATCGAGATGCCAAGCTGGAAATGCTCGCACTC	3234
Qy	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu	80
Db	3235	CTCTTGATGATGATTATGAGAGCATCGCTGGCCCTGGGCTAGCTCCATTAGCATCTTTGAA	3294
Qy	81	AsnMetLeuLeuArgThrLeuSerGluLysAlaArgAspMetLysLys-----	97
Db	3295	AACATGAACCTGCGAACCTCTCGAGAAAGGCACGGATGACATGAAGAAGACATTCACCA	3354
Qy	97	-----	97
Db	3355	GAAGATCTCTGAAGCAACGATGACTGACCACGACCAAGCAAGAAAGTCCAGGAATT	3414
Qy	98	-----IleSer	99
Db	3415	TCACAAGCTGTGCAACAGATAGTGCCACAGCTGCAGAGACAAAGAACAAGAAATTTCA	3474
Qy	100	GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlnIleSerGln	119
Db	3475	CAAGCTATGGAACAAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGAAATTTCAAA	3534
Qy	120	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGlyAsp	139
Db	3535	GCTATGGACAAAGAGGTGCCACAGCAGCAGCAGACAGACAGACAGACAGACAGCATGGAGGTGAC	3594
Qy	140	ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg	159
Db	3595	ACATGGGACTCAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGT	3654
Qy	160	SerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp	179
Db	3655	AGTTTTGAAACACATGCTGCTGACCTGGCCGGAATTCGCAACGTTGGCTGGTCTTTGAT	3714
Qy	180	SerAspArgTgPhePheArgProArgThrValValLeuHisGlyLysSerGlyIleGly	199
Db	3715	TCAGACCGGTGGGCTTCCGGCCCTCGACCGTGGTTCGACACGGAAGTTCAGGAATTTGGG	3774
Qy	200	LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGly	219

Db	3775	AAATCGGCTCTAGCCAGAGGATCGTCTGTCTGGCGCGCAAGGTGAGTCTTACCAGGGA	3834
Qy	220	MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer	239
Db	3835	ATGTTCTTCTTACGTCTTCTTCTCCCGTTAGAGAGATGACGCGGAAGAAGAGAGCAGT	3894
Qy	240	ValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet	259
Db	3895	GTCAAGAGTTCTATCTCCAGGAGTGCCAGACTCCCGAGGCTCCGCTGACGGAGATCATG	3954
Qy	260	SerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySerValLeu	279
Db	3955	TCCCGACCAAGAAAGGCTGTTTTCATCATTCAGCGGTTTCGATGACCTGGGCTCTGTCTC	4014
Qy	280	AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle	299
Db	4015	AACAATGACAAAGCTCTGCAAGAGCTGGCTGAGNAGAGGCTCGTTTCAACCTCATTA	4074
Qy	300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp	319
Db	4075	CGCAGTCTGCTGAGGAAGTCT	4134
Qy	320	ValGlyThrGluLysLysSerGluValValSerProArgTyrLeuLeuValArgGly	339
Db	4135	GTGGGCACAGAGAGCTCAAGTCAGAGGCTGTCTCTCCCGTTTACCTGTTAGTAGGGA	4194
Qy	340	IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys	359
Db	4195	ATCTCCGGGGAACAAGAATCCACTTGTCTCTTGGCGCGGATTTGGTAGCATCAGNAG	4254
Qy	360	ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro	379
Db	4255	ACACAAGGTTGCTGCGATCATGAACAACCGTGAGCTGCTCGACCAAGTCCAGGTGCC	4314
Qy	380	AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal	399
Db	4315	CGCGTGGGCTCTCTCATCTGCGTGGGCTGACGCTGAGACGCTGGTGGGGAGAGGCTC	4374
Qy	400	AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr	419
Db	4375	GCCTCTTCAACCAACGCTCACAGGCTGCACGCGCTTTTGTGTTTTCATCAGCTCACC	4434
Qy	420	ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe	439
Db	4435	CCTCGAGGCTGCTGCGGCTGCTCAATCTCGAGGAAAGAGTTGCTCGTAAGCGCTTC	4494
Qy	440	CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu	459
Db	4495	TGCCGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGACGGTGAAGCCTC	4554
Qy	460	MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu	479
Db	4555	ATGTTTCAAGACTCGGGAGTCTGAGCTCCGTCTCTGTTTTCATGAACATCTCTTCTC	4614
Qy	480	ProAspSerHisCysGluGlyTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys	499
Db	4615	CCAGACGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGGACTTCTGT	4674
Qy	500	AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr	519
Db	4675	GCCTCTTGTACTACGTGTTAGAGGCGCTGGAATTCGAGCAGCTCTCTGCTCTCTCTCT	4734
Qy	520	ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu	539
Db	4735	GTTGAGAAGCAAGAAGGCTCATGGAGCTTAAACAGCAGGCTTCCATATATCCACTCCG	4794
Qy	540	TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal	559
Db	4795	TGATGAGGCTTCTTGTGCTCTGAGCGAGACGCTAGAGGAGCCACCTGGAGGCTC	4854
Qy	560	LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu	579
Db	4855	CTGCTGGGCTGTCCCGTTCCCTGGGGTGAAGCAGAAGCTTCTGCACTGGGCTCTCTCTG	4914

```
QY 580 LeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTGAGCAGCTTAATGCCACACCCAGAGACACCCCTGGACGCTTCCACTGTCCT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAAACAGCTTCCAAAGAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCCGATTAACCAAGAACCTCGACTTGATAGCATCTTCTCTCCCTCCAGCACTGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTCGGGAANAATCGGCTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCCTGTGGTCCCTCTATGATGCGGATAAGACCCCTCATTTGAGGAGCAGTGGAA 5214
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 5215 GATTTCGTCTCATGCTTGGCACCACCCACACCTGGCGGAGCTGGACCTGGGCGAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
DB 5275 ATCTGACAGACGGGGCATGAAGACCTGTGTCCCAAGCTGAGGCATCCACCTGCAAG 5334
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
DB 5335 ATACAGACCCCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGACACCTCTGGAGA 5394
QY 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
DB 5395 ATCGTCATGGCAACCGCTAAGATCCCTCAACTTGGGAGGCACCCACCTCGAAGGAA 5454
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuGluSerLeu 779
DB 5455 GAGGATGTAGGATGGGTGTGAAGCCCTTAACACCCCAAAATGTTTGTGGAGTCTTTG 5514
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
DB 5515 AGGCTGATGCTGTGGATTGACCCATGCTGTACCTGAAGATCTCCCAAACTCTTACG 5574
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
DB 5575 ACCTCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAGGGAGTA 5634
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
DB 5635 ATGCCTCTCAGTGATGCTTGAGAGTCTCCAGTGCCTGCGCCCTGCAAGAGCTGATCTGGAG 5694
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
DB 5695 GACTGTGGCATCACAGCACCGGTGGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 5754
QY 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
DB 5755 AGCTTGACACACCTGTGCTATCCAAACAACAGCCCTGGGGAAACGAAGGTGTAATCTACTG 5814
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
DB 5815 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGGCCAC 5874
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
DB 5875 CTGGACACGGCTGCTGGTGTTCCTTGCACTTGCCTTATGGGTAACTCATGGCTGACG 5934
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
DB 5935 CACCTGAGCCTTAGCATGAACCTGTGGAAACAATGGCGTGAAGCTTCTGTGCGAGGTC 5994
```

```
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
DB 5995 ATGAGAGAACCATCTTTGTATCTCCAGGACCTGGAGTTGTAAGTGTATCTATCTACCGCC 6054
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
DB 6055 GCGTGTCTGTGAGAGTCTGTCTCTGTGTGATCTCGAGGAGCAGACACCTGAAGACCTGGAT 6114
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
DB 6115 CTCACGACAAATGCCCTGGGTGACGGTGGGGTGTGTCACCTGTGCGAGGACCTGAAGCAA 6174
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
DB 6175 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTGCTGT 6234
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
DB 6235 GAGGCACCTCTCTTGGCCCTTTCTCTGCAACCGGCATCTGACCAGTCTAAACCTGGTGCAG 6294
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
DB 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGCGCCTTTGCTGTCCTCCACGTCT 6354
QY 1060 AsnLeuGlnIleGly----- 1065
DB 6355 AACTTACAGATAATTGGCAATGACTCTGAAGAAATGACGTTCTTCGAGAATCTGCTCTA 6414
QY 1065 ----- 1065
DB 6415 GTAGTTTGTCTTAAAGTCACTGTTTCCAAAGAACCTATCAATGACATTAAGGGAGAACTTA 6474
QY 1065 ----- 1065
DB 6475 CTGTACTCTCCCAACCCCTATAACAGACCCGSCACAGAGACTCTCGAGGAAGAAATTCAT 6534
QY 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
DB 6535 GGTGAGCTGAAAGGCTGTGGAATGGCAGTACCTCTGCAATAATAGGAAGCTGCTGGAG 6594
QY 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
DB 6595 GAAGTCAGCTACTCAAGCCCGAGTCGTAATTTGACGGTAGTTGGCATCTCTTTTGATGAA 6654
QY 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
DB 6655 GATGACCGAATTGGATCTTCAAAGCCCAAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
QY 1112 LeuProGluSerArg 1116
DB 6715 CTCATGGCATCAG 6729
RESULT 9
ADA45220
ID ADA45220 standard; DNA; 3830 BP.
XX ADA45220;
XX
XX 20-NOV-2003 (first entry)
XX Human MATER splice variant DNA SEQ ID 3.
XX human; MATER; maternal antigen that embryos require; NTP-ase;
XX nucleoside triphosphate-ase; apoptosis; antifertility; contraceptive;
XX antiinflammatory; immunosuppressive; gynaecological; endometriosis;
XX ovarian dysfunction; autoimmune premature ovarian failure;
XX autoimmune disease; contraceptive; autoantibody; ds; gene.
XX Homo sapiens.
XX Key
XX Location/Qualifiers
XX 1..3432
XX /*tag= a
FT
```

/product= "splice variant of MATER"

FT EP1285964-A2.
PN 26-FEB-2003.
XX 12-JUL-2002; 2002EP-00090246.
PD 10-AUG-2001; 2001DE-01039874.
PF (SCHD) SCHERING AG.
PR Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
PA WPI; 2003-302814/30.
XX P-PSDB; ADA45221.
XX New nucleic acid for human maternal antigen that embryos require protein,
PT useful e.g. in diagnosis and treatment of female infertility.
PS Claim 1; Page 17-18; 31pp; German.
XX This invention describes a novel human MATER (maternal antigen that
CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
CC associated with apoptosis, a defect in MATER activity causes growth
CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
CC MATER gene is found predominantly in ovary, testis and placenta. The
CC products of the invention have anti-infinity, contraceptive,
CC anti-inflammatory, immunosuppressive and gynaecological activity.
CC Effectors of the MATER polypeptide are used to treat infertility.
CC associated with endometriosis and also ovarian dysfunction, autoimmune
CC premature ovarian failure, inflammation, autoimmune diseases and female
CC infertility, and as contraceptives. Measurements of autoantibodies
CC against MATER in body samples or MATER protein or mRNA in eggs is useful
CC for diagnosis of female infertility, particularly by detecting mutations
CC in the MATER gene with a DNA chip. This sequence represents a splice
CC variant of human MATER in which exon 4 of the usual form is absent.
SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3830
Score: 5683.50 Matches: 1087
Percent Similarity: 95.87% Conservative: 5
Best Local Similarity: 95.43% Mismatches: 12
Query Match: 80.34% Indels: 35
DB: 7 Gaps: 2
US-10-066-521-6 (1-1344) x ADA45220 (1-3830)
QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
DB 1 ATGGAAGAGAGACAAATCGCTCCACCTTTCCAGCTACGGCTGCAATGGTGCTCTATGAG 60
QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
DB 61 CTAGACAAGGAGAAATTCAGACATTCAGGAATTAAGAAAGAAATCTTCAGAAATCG 120
QY 41 ThrThrCysSerIleProGlnPheGluLeuGlnAenAlaAenValGluCysLeuAlaLeu 60
DB 121 ACCACATGCTCTATTCCACAGTTTGAATTCGAGAAATGCCAAGTGGAAATGTCTGCACCT 180
QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTpaAlaThrSerIleSerIlePheGlu 80
DB 181 CTCCTTGCATGATATTATGAGCATCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 240
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGln 100
DB 241 AACATGAACCTCGGAACCTCTCGGAGAGAGGACCGGATGACATGAAATTCACCAGAA 300
QY 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln--- 115
DB 301 GATCCTGAAGCAACGATGACTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360

QY 115 ----- 115
DB 361 TATGGCATGACTAAGCTTATCTTTGGGGGTCTCTGACATCTCTGACTCGAATAATAAACAC 420
QY 116 -----GluIleSerGlnAlaMetGluGlnGluGly 125
DB 421 AAGTATGTTGGAATTCATTCTTTTTCAGAAATTTTCAAGCTATGGAACCAAGAGGT 480
QY 126 AlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
DB 481 GCCACAGCAGCAGACAGACAGAAAGACACAGGACATGGAGTGCACATGGGACTACAGAGT 540
QY 146 HisValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAenThrAla 165
DB 541 CACGTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTCTGTGTTTGAATTCAGACCCGGTGGGCTTC 600
QY 166 AlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPhe 185
DB 601 GCTGACTGGCCGGAATTCGAAACGTTGGCTGGTCTTTTGAATTCAGACCCGGTGGGCTTC 660
QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 205
DB 661 CGGCTCGCACGGTGGTCTTCGACGGAAGTCAAGAAATTCGGAATTCGGCTCTAGCCAGA 720
QY 206 ArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225
DB 721 AGGATCGTGTCTGGTGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCTCTAGCTCTTC 780
QY 226 PheLeuProValArgGluMetGlnArgLysGlyLysSerValThrGluPheIleSer 245
DB 781 TTCTCTCCCTTTAGAGAGATGTCAGCGGAAGAGAGAGAGATGTCACAGAGTTTCATCTCC 840
QY 246 ArgGluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeu 265
DB 841 AGGAGTGGCCACACTCCAGGCTCCGGTACGAGATCATGTCCCGACCAAGAGCTG 900
QY 266 LeuPheIleAspGlyPheAspAspLeuGlySerValLeuAenAenAspThrLysLeu 285
DB 901 TTGTTCATCATTCAGCGGTTTCGATGACCTGGGCTCTGTCTCAACATGACACAAAGCTC 960
QY 286 CysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeuArgLys 305
DB 961 TGCAAGACTGGGCTGAGAAAGACGCTCCGTTCCACCTCATACGACGTCTGCTGAGGAAG 1020
QY 306 ValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeu 325
DB 1021 GTCTGTCTCTCTGAGTCTTCTCTGATCGTCCGCTCAGAGACGTGGGCACAGAGAGCTC 1080
QY 326 LysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArg 345
DB 1081 AAGTCAGAGTCTGTCTCCCGTTACCTGTAGTAGAGGAATCTCCGGGGAACAAAGA 1140
QY 346 IleHisIleLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAla 365
DB 1141 ATCCACTTGTCTCTGAGCGGGGATTTGGTGAGCATCAGAAGACACAAAGGTTGCGTGC 1200
QY 366 IleMetAenAenArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIle 385
DB 1201 ATCATGAACACCGTGAGTGTCTCGACAGTGCAGGTCGCGGCTGGGCTCTCTCATC 1260
QY 386 CysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThr 405
DB 1261 TGCCTGGCCCTGACGTGTCAGGAGCTGTGGGGGAGAGCGTCCGCCCTTCAACCAACG 1320
QY 406 LeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArg 425
DB 1321 CTCACAGGCTGCACGCGGCTTTTGTGTTTCATCAGCTCACCCCTCAGGCGGTGGTCCGG 1380
QY 426 ArgCysLeuAenLeuGluArgValValLeuLysArgPheCysArgMetAlaValGlu 445
DB 1381 CGCTGTCTCAATCTGGAGGAAGAGTTGTCTCTGAAAGCGCTCTCTGCGGTATGGCTGTGGAG 1440

QY 446 GlyValTyrAsnArgLysSerValPheAspGlyAspLeuMetValGlnGlyLeuGly 465
DB 1441 GGAGTGTGGAAATAGGAAGTCTAGTGTGGACGGTGACGACCTCAATGGTTCAAGGACTCGGG 1500
QY 466 GluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGlu 485
DB 1501 GAGTCTGAGCTCCGTGCTCTGTTTACATGAACATCTCTTCCACAGACGCCACTGTGAG 1560
QY 486 GluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrVal 505
DB 1561 GAGTACTACACCTTCTCCACCTCAGTCTCCAGGACTTCTGTGCGGCTTGTACTAGGTG 1620
QY 506 LeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLeuYsThrLysArg 525
DB 1621 TTAGAGGCGCTTGGAAATCGAGCCAGCTCTCTGCGCCCTCTGTACGTTGAGAAGACAAAGG 1680
QY 526 SerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyrMetLysArgPheLeu 545
DB 1681 TCCATGGAGCTTAAACAGCGAGGCTTCCATATCCACTCGCTTTCGATGAAGCGTTCTTG 1740
QY 546 PheGlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProVal 565
DB 1741 TTTGGCCTCGTGAGCGAAGAGCTTAAGAGGCGCACTGGAGGTCCTGCTGGGCTGTCCGTT 1800
QY 566 ProLeuGlyValLysGlnLysLeuLeuHisTyrValSerLeuLeuGlyGlnGlnProAsn 585
DB 1801 CCCCTGGGGGTGAAGCAGACCTTCTGCACCTGGGTCTCTGTGGGTGACGACCTAAT 1860
QY 586 AlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLys 605
DB 1861 GCCACCACCCAGAGAGACACCTCGAGCGCTTCCACTGCTCTTTCGAGACTCAAGACAAA 1920
QY 606 GluPheValArgLeuAlaLeuAsnSerPheGlnGluValTyrLeuProIleAsnGlnAsn 625
DB 1921 GAGTTTGTTCGCTTGGCATTAAACAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAAC 1980
QY 626 LeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArg 645
DB 1981 CTGACCTTGATAGCATCTTCTTCTGCTCCAGCAGCTGCTCGTATTTGCGGAAATTCGG 2040
QY 646 ValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValPro 665
DB 2041 GTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGCAATGTCCTGTGTCCT 2100
QY 666 LeuTyrMetArgAspLysThrLeuIleGluGlnTyrGluAspPheCysSerMetLeu 685
DB 2101 CTATGGATGCGGGATAAGACCTCATTTGAGAGCAGTGGGAAGATTCTTGTCTCCATGCTT 2160
QY 686 GlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAla 705
DB 2161 GGCACCCACACACCTGCGGACGTGGACCTGGGACAGCAGCATCCCTGACAGACGGGGCC 2220
QY 706 MetLysThrLeuCysAlaLeuLysArgHisProThrCysLysIleGlnThrLeuMetPhe 725
DB 2221 ATGAAGACCTGTGTGCCAAGCTGAGGCATCCCACTGCAAGATACAGACCTCATGTTT 2280
QY 726 ArgAsnAlaGlnIleThrProGlyValGlnHisLeuTyrArgIleValMetAlaAsnArg 745
DB 2281 AGAAATGCACAGATTACCCCTGGTGTGCAGACACCTCTGGGAATCGTATGGCCCAACCGT 2340
QY 746 AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAla 765
DB 2341 AACCTAGATCCCTCACTTGGGAGGCACCCACTGGAAGGAGGATGTAAAGATGGCG 2400
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGly 785
DB 2401 TGTGAAGCCTTAAACACCCAAAATGTTTGTGGAGTCTTTGAGGCTTGGATTGTGTGGA 2460
QY 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
DB 2461 TTGACCATGCTGTGTTACCTGAAAGATCTCCAAAATCTCTTACGACTCCCCAGCCTGAAA 2520
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825

DB 2521 TCTCTGAGCCTGCAGGAAACAGGTCAGACACCGGAGTAATGCTCTCAGTGATGCC 2580
QY 826 LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAla 845
DB 2581 TTGAGAGTCTCCAGTGGCGCTTCAGAAAGCTGATACTGGAGGACTGTGGCATCAGACC 2640
QY 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
DB 2641 ACCGGTTCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGC 2700
QY 866 LeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeu 885
DB 2701 CTATCCAAACACAGCCTGGGAAACGAAGGTGTAAATCTACTGTGTCATCCATGAGGCTT 2760
QY 886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
DB 2761 CCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCCACCTGGACACGGCTGGCTGT 2820
QY 906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTyrLeuThrHisLeuSerLeuSerMet 925
DB 2821 GGTTCCTTTCGACTTGGCTTATGGGTAACTCATGCTGACGACCTGAGCCTTAGCATG 2880
QY 926 AsnProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCys 945
DB 2881 AACCTGTGGAGACAAATGGCGTGAAGCTTCTGTGCGAGTCTATGAGAGACCATTTGT 2940
QY 946 HisLeuGlnAspLeuGluValLysCysHisLeuThrAlaAlaCysCysGluSerLeu 965
DB 2941 CATCTCCAGACCTGGAGTTGGTAAAGTGTCTCATCTCACCGCGCTGTGTGAGAGTCTG 3000
QY 966 SerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeu 985
DB 3001 TCTGTGTGATCTCGAGGACAGACACCTGAAGAGCTTGATCTCAGGACAAATGCCCTG 3060
QY 986 GlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThr 1005
DB 3061 GGTGACGGTGGGTTTCTGCACCTGTGCGAGGACTGAGCAAAAAGAACAGTGTCTGACG 3120
QY 1006 ArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAla 1025
DB 3121 AGACTCGGGTTGAAGCATGTGGACTGACTTCTGATTGCTGTGAGGCACTCTCTTGGCC 3180
QY 1026 LeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLys 1045
DB 3181 CTTTCTGTCAACCGGCATCTGACAGTCTAAACCTGGTGAGAAATAACTTCAGTCCCAAA 3240
QY 1046 GlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeGly 1065
DB 3241 GGAATGATGAGCTGTGTGGCCTTTGCTGCTCCACGCTTAACCTTACAGATAATTGGG 3300
QY 1066 LeuTyrLysTyrPdnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeu 1085
DB 3301 CTGTGAAAATGGCAGTACCTGTGCAAAATAAGAAAGCTGTGAGGAAGTGCAGCTACTC 3360
QY 1086 LysProArgValIleLeuAspGlySerTyrHisSerPheAspGluAspArgHis 1104
DB 3361 AAGCCCCGAGTGTAAATGACGGTAGTGTGCATCTTTTGTATGAAGATGACCGGTAC 3417
RESULT 10
ABX97181
ID ABX97181 standard; cDNA; 3226 BP.
XX
AC ABX97181;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV125b cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
human; gene; ss.
XX


```
OS Homo sapiens.
XX WO200272757-A2.
XX 19-SEP-2002.
XX 08-MAR-2002; 2002WO-US006908.
XX 08-MAR-2001; 2001US-0274101P.
XX 08-MAR-2001; 2001US-0274194P.
XX 08-MAR-2001; 2001US-0274281P.
XX 08-MAR-2001; 2001US-0274322P.
XX 09-MAR-2001; 2001US-0274849P.
XX 12-MAR-2001; 2001US-0275235P.
XX 13-MAR-2001; 2001US-0275578P.
XX 13-MAR-2001; 2001US-0275579P.
XX 13-MAR-2001; 2001US-0275601P.
XX 14-MAR-2001; 2001US-0276000P.
XX 16-MAR-2001; 2001US-0276776P.
XX 19-MAR-2001; 2001US-0276994P.
XX 20-MAR-2001; 2001US-0277239P.
XX 20-MAR-2001; 2001US-0277321P.
XX 21-MAR-2001; 2001US-0277327P.
XX 21-MAR-2001; 2001US-0277791P.
XX 22-MAR-2001; 2001US-0277833P.
XX 23-MAR-2001; 2001US-0278152P.
XX 26-MAR-2001; 2001US-0278894P.
XX 27-MAR-2001; 2001US-0278999P.
XX 27-MAR-2001; 2001US-0279036P.
XX 28-MAR-2001; 2001US-0279344P.
XX 30-MAR-2001; 2001US-0277338P.
XX 30-MAR-2001; 2001US-0279995P.
XX 30-MAR-2001; 2001US-0280233P.
XX 02-APR-2001; 2001US-0280802P.
XX 02-APR-2001; 2001US-0280822P.
XX 02-APR-2001; 2001US-0280900P.
XX 04-APR-2001; 2001US-0281194P.
XX 13-APR-2001; 2001US-0283675P.
XX 30-APR-2001; 2001US-0287424P.
XX 02-MAY-2001; 2001US-0288066P.
XX 03-MAY-2001; 2001US-0288342P.
XX 03-MAY-2001; 2001US-0288528P.
XX 15-MAY-2001; 2001US-0291190P.
XX 16-MAY-2001; 2001US-0291099P.
XX 16-MAY-2001; 2001US-0291240P.
XX 30-MAY-2001; 2001US-0294485P.
XX 31-MAY-2001; 2001US-0294889P.
XX 31-MAY-2001; 2001US-0294899P.
XX 18-JUN-2001; 2001US-0299027P.
XX 19-JUN-2001; 2001US-0299303P.
XX 10-JUL-2001; 2001US-0304354P.
XX 31-JUL-2001; 2001US-0309198P.
XX 16-AUG-2001; 2001US-0312903P.
XX 10-SEP-2001; 2001US-0318462P.
XX 12-SEP-2001; 2001US-0318770P.
XX 27-SEP-2001; 2001US-0325430P.
XX 27-SEP-2001; 2001US-0325681P.
XX 18-OCT-2001; 2001US-0330380P.
XX 31-OCT-2001; 2001US-0335301P.
XX 14-NOV-2001; 2001US-0332172P.
XX 14-NOV-2001; 2001US-0332271P.
XX 14-NOV-2001; 2001US-0332272P.
XX 14-NOV-2001; 2001US-0333184P.
XX 14-NOV-2001; 2001US-0333272P.
XX 21-NOV-2001; 2001US-0332094P.
XX 03-DEC-2001; 2001US-0337426P.
XX 03-DEC-2001; 2001US-0338092P.
XX 04-DEC-2001; 2001US-0337185P.
XX 03-JAN-2002; 2002US-0345705P.
XX 07-MAR-2002; 2002US-00092900.
XX PA (CURA-) CURAGEN CORP.

XX PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
XX P-PSDB; ABU65214.
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
a disorder associated with aberrant NOVX expression or activity e.g.,
cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
asthma.
XX Claim 13; Page 467; 1103pp; English.
XX This invention describes novel human NOVX polypeptides which have
cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
activity. Pharmaceutical compositions comprising the NOVX proteins or
nucleic acid molecules or NOVX antibodies are useful for preventing or
treating a disorder associated with aberrant NOVX expression or activity
e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
asthma. The products of the invention can be used for gene therapy or in
a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
ABU65041-ABU65218
XX SQ Sequence 3226 BP; 738 A; 856 C; 914 G; 718 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3226
Score: 5120.00 Matches: 977
Percent Similarity: 96.95% Conservative: 7
Best Local Similarity: 96.26% Mismatches: 20
Query Match: 72.38% Indels: 11
DB: Gaps: 1

US-10-066-521-6 (1-1344) x ABX97181 (1-3226)
Qy 101 AlaMetGluGluGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAla 120
Db 113 GCCCAAGGGAGGAGGCTCAGGCCAGCGAGACGCGACGCGAAGAAACGTCACGAAAAA 172
Qy 121 MetGluGluGluGlyAlaThrAlaAlaGluThrGluGluGln-Gly----- 135
Db 173 CAGCCTCAGTTGCCAGTCCCTTTCAGGAAACAGACAGGCGCTGGGGCGGCCACCTGGG 232
Qy 136 -----HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLy 150
Db 233 CTCAGAGCTTGGGCTGTCATGGAGGTGACATGGGACTACAAGAGTCACGTGATGACCA 292
Qy 150 sPheAlaGluGluGluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGl 170
Db 293 ATTCCGCTGAGGAGGAGATGCTAGCTGCTAGTTTGAACACACTGCTGCTACCTGGCCGA 352
Qy 170 uMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrVa 190
Db 353 AATGCAAAACGTTGGCTGCTTTTTCATTTCAGACCGGCTGGGGCTTCGGGCTCGCAGGT 412
Qy 190 lValLeuHisGlyLysSerGlyLeGlyLysSerAlaLeuAlaArgGlyLeuValLeuCy 210
Db 413 GGTTCGTCACGGAAGTCAGGAATTGGAAATCGGCTCTAGCCAGAGGATCGTGTGTG 472
Qy 210 sTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValAr 230
Db 473 CTGGGCGCAAGGTGGACTCTACCCAGGGAATGTTCTCTCTAGCTCTTCTCTCCCGTTAG 532
Qy 230 gGluMetGlnArgLysLysGluSerValThrGluPheIleSerArgGluTrpProAs 250
Db 533 AGAGATGCGCGGAAGAGGAGGAGCAGTGTCAACAGAGTTTCATCTCCAGGAGTGGCCAGA 592
```

QY 250 pSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAs 270
Db 593 CTCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGA 652
QY 270 pGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrrAl 290
Db 653 CGGTTTCGATGACCTGGGCTCTGTCTCTCAACATGACACAAGCTCTGCAAGACTGGC 712
QY 290 aGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGl 310
Db 713 TGAGAAGCAGCGCTCCGTTACCCCTCATACGAGTCTGCTGAGGAAGTCTCTGTCTCCCTGA 772
QY 310 uSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValva 330
Db 773 GTCCCTTCTGATCGTCAACGTCAGAGCTGGGCACAGAGAGCTCAAGTCAGAGGTCGT 832
QY 330 lSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLe 350
Db 833 GTCTCCCGTTACCTGTAGTTAGAGNAATCTCCGGGGAACAAGAAATCCACTTGTCTCT 892
QY 350 uGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAr 370
Db 893 TGAGCGCGGATTTGGTGAGCATCAGAAGACACAAGGGTTGGTGGCATCATGAACAACCG 952
QY 370 gGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGl 390
Db 953 TGAGTGTCTGACCAAGTCCAGGTGCGCGCTGCTCTCATCTGCGTGGCCCTGCA 1012
QY 390 nLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHi 410
Db 1013 CCTGCAGGACGTGTGGGGAGAGGCTGCGCCCTTCAACCAACGCTCACAGGCTTGA 1072
QY 410 aAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLe 430
Db 1073 CGCGCTTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCT 1132
QY 430 uGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnAr 450
Db 1133 GGAGAAAGAGTTGTCTGAAGCGCTTCTGCGGTATGGCTGTGAGGAGGTGGGAATG 1192
QY 450 GlySerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuAr 470
Db 1193 GAAGTCAGTGTTCAGCGTGACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCG 1252
QY 470 gAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPh 490
Db 1253 TGCTCTGTTTCACATGAACATCTCTTCCACAGACGACCTGTGAGGAGTACTACACTT 1312
QY 490 ePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGl 510
Db 1313 CTTCCACCTCAGTCTCCAGGACTTCTGTGCGGCTTGTACTACGTGTAGAGGCGCTGGA 1372
QY 510 uIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLy 530
Db 1373 AATCGAGCCAGCTCTCTGCGCTCTGTAGCTTGAGAGACAAGAGGTCCATGGAGCTTAA 1432
QY 530 sGlnAlaGlyPheHisIleHisSerLeuTrrpMetLysArgPheLeuPheGlyLeuValse 550
Db 1433 ACAGCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTGGCTCGTGAG 1492
QY 550 rGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVally 570
Db 1493 CGAAGACGTAAAGAGGCGCACTGGAGGTCTCTGCGGCTGTCCTGCTCCCTGGGGGTGA 1552
QY 570 sGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGl 590
Db 1553 GCAGAGCTTCTGACTGGGTCTCTCTGTTGGGTTCAGCAGGCTTAATGCCACCCACCG 1612
QY 590 yAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLe 610
Db 1613 AGACACCTGGACGCTTCCACTGTCTTTTCGAGACTCAAGACAAGAGTTTGTTCGCTT 1672

QY 610 uAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAl 630
Db 1673 GGCAATTAACAGCTTCCAAGAGTGTGGCTTCGATTAACAGAACCTGAGACTGTATAGC 1732
QY 630 aSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspVallyLysGl 650
Db 1733 ATCTCTCTCTGCTCCAGCACTGTCCGATATTTCGGGAAAATTCGGGTGGATGATCAAAG 1792
QY 650 yIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetArgAs 670
Db 1793 GATCTTCCCAAGAGATGAGTCCGCTGAGGCATGCTCCTGTGTGCTCCCTCATATGGATGGGA 1852
QY 670 pLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHi 690
Db 1853 TAAGACCTTCATTGAGAGAGCAGTGGGAAGATTCTCTCATCTTGGCACCCACCACCA 1912
QY 690 sLeuArgGlnLeuAspLeuGlySerIleLeuThrGluArgAlaMetLysThrLeuCy 710
Db 1913 CCTGCGGCACTGACCTGGGCAGCAGCATCTCTGACAGAGCGGCCCATGAAGACCTGTG 1972
QY 710 sAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIl 730
Db 1973 TGCCAAAGCTGAGCATCCCACTGCAAGATACAGACCTGATGTTTAGAAATGCACAGAT 2032
QY 730 eThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLe 750
Db 2033 TACCCCTGGTGTGCAGACCTCTGGAGAAATCGTCATGGCCCAACCGTAACCTAAGATCCCT 2092
QY 750 uAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLy 770
Db 2093 CAACTTGGGAGGCCACCCACTGAAGAGAGGATGTAAGATGGCGTGTGAACCTTAA 2152
QY 770 sHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCy 790
Db 2153 ACACCCAAAATGTTCTTGGAGTCTTTGAGGCTGGATGCTGTGGATGACCCATGCTG 2212
QY 790 sTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAl 810
Db 2213 TTACCTGAGAGATCTCCAAATCTTACGACCTCCCCAGCCTGAAATCTCTGAGCCTGGC 2272
QY 810 aGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGl 830
Db 2273 AGGAAACAAGTGACAGACCCAGGAGTAATGCCTCTCAGTGATGCCTTGAGAGTCTCCCA 2332
QY 830 nCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSe 850
Db 2333 GTGCGCCTCGCAGAGCTGATACTGAGGAGCTGTGGCATCACAGCCACGCGTTGCCAGAG 2392
QY 850 rLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSe 870
Db 2393 TCTGGCCTCAGCCTCGTCAGCAACCGAGCTTGACACACCTGTGCTATCTCCAACACAG 2452
QY 870 rLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLe 890
Db 2453 CCTGGGGAACGAGGTGTAATCTACTGTGTGATCCATGAGGCTTCCCACCTGTAGTCT 2512
QY 890 uGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLe 910
Db 2513 GCAGAGGCTGATGCTGAAATCAGTCCACCTGGACACGCGTGGCTGTGGTTTTCTTGCACT 2572
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 2573 TGGCCTTATGGGTAACTCATGGCTGACGACCTTGAGCCTTAGCATGAACCTGTGAAGA 2632
QY 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 2633 CAATGGCGTGAAGTCTCTGTGCGAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCT 2692
QY 950 uGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 2693 GGAGTTGGTAAAGTGTCACTCTCACCCCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 2752
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990

```

Db 2753 GAGGAGCAGACACCTGAAGAGCCTGGATCTACGACCAATGCCCTGGGTGACGGTGGGGT 2812
Qy 990 lAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
Db 2813 TCCTGCATGTGCGAGGAGTGAAGCAAAAGACAGTGTCTGACGACACTCGGGTTGAA 2872
Qy 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 2873 GGCATGTGACCTGACTTCTGATTGCTGTGAGGCACTCTCCTTGGCCCTTCTCTGCAACCG 2932
Qy 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 2933 GCATCTGACCATCTAAACCTGGTGGTGAATAAATTCAGTCCCAAGGAATGATGAAGCT 2992
Qy 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnLeuGlyLeuThrLysTrpG1 1070
Db 2993 GGTTCGGCCCTTGGCTGTGCCACGCTTAACCTACAGATAATGGGCTGTGGAATGGCA 3052
Qy 1070 nTyrProValGlnIleArgLysLeuGluValGlnLeuLysProArgValVa 1090
Db 3053 GTACCTGTGCAATAAAGGAAGCTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGGT 3112
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3113 AATTGACGGTAGTGGCATCTTTTGTATGAAGATGACCGGTAC 3155

RESULT 11
ABK48611
ID ABK48611 standard; cDNA; 3447 BP.
XX
AC ABK48611;
XX
DT 13-AUG-2002 (first entry)
XX
DE Mouse MATER cDNA.
XX
KW Mouse; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; WATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; gene therapy; fertility.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT 9..3344
FT CDS /tag= a
FT /product= "Mouse MATER"
XX
XX WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR P-PSDB; AAU79525.
XX
XX New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
XX Example 5; Page 72-74; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a WATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell

```

```

CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene. The MATER protein and
CC polynucleotide, by gene therapy, are useful in diagnosing or treating
CC fertility and reduced fertility. In particular, the MATER protein is
CC useful as a contraceptive agent, or for influencing (either inhibiting or
CC enhancing) fertility and can be used to detect a predisposition to
CC infertility or reduced fertility, or for presymptomatic screening of an
CC individual for infertility/reduced fertility. The protein and
CC polynucleotide are also useful for detecting an excess or deficiency, or
CC genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the mouse MATER cDNA
XX
SQ Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 2.94e-204 Length: 3447
Score: 2688.00 Matches: 554
Percent Similarity: 66.49% Conservative: 186
Best Local Similarity: 49.78% Mismatches: 293
Query Match: 38.00% Indels: 81
DB: 6 Gaps: 11

US-10-066-521-6 (1-1344) x ABK48611 (1-3447)
Qy 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPheLysGluLeuLys 34
Db 163 CAATGGGTCTCCCA-GAAAAAGACAGTAAGCAATCTTGAAGCACCCTGGATTGGAAGAG 221
Qy 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
Db 222 GAACAGAAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
Db 243 ---ATGTCTCTTCAGAAATATGTCAGTAGAGCAATCTCTGAAGACAGTGGATCAAGAGAA 299
Qy 95 MetLysLysIleSerGln----- 100
Db 300 GTGGAACACGGCTCAGAAAGAAAAATGACTTCTCCAGAAAAACGACGATAAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAAGACCAAGGACCAGAGCAGAGCAGACATCAGAAACCTTACAACTTAAGGAAGAAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTGACAGAGGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 444 AATGGAGGTGACTTACAGACTACAGGCCCATGTGATTGCTAAGTTGCGACACAAGTGTG 503
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCGACAGATGAAATTTATGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATCCTTTTAAACCATACAGAAACCTTCCAGCTTCACCATATATCCATATCCATGGAAGA 602
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 603 CCAGAGTGTGGGAAGTCAAGCTTTGGCCAGAAATATTGTTCTTGGCTGGGCACAGGGTAA 662
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235

```



```
QY 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTATGGGACTGCTCACTCACAGAACTGCTCGGAGGACCTGGCTGTATGATCAACA 2930
QY 972 SerArgHisLeuSerLeuAspLeuThrAspAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAGACCTTAAAGATTGATCTTGGTAACACGCCCTGGGTGACAAAGGAGTCATA 2990
QY 992 AlaLeuCysGluGlyLeuLeuGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAACGAAGTAGCAGCTCCCTGAGGAGACTTGGGTGGGGCA 3050
QY 1012 CysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAATTGACTTCCAATTGCTGTGAGGCATTGTCAATTGGCCATCTCTTGAACCCCTCAC 3110
QY 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLeuLysCys 1051
Db 3111 CTGAACAGCCTAAACCTGTGAGAGATGACTTCAGTACATCGGGATGTTGAAGCTGTGC 3170
QY 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrp 1071
Db 3171 TCTGGTTTCCAATGCCCTGTCTTAACCTGGGGATAATTGCTGTGGAACGAGGTAC 3230
QY 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValValle 1091
Db 3231 TATGCCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTTGTCAAGCCGCCACGTGCTGATT 3290
QY 1092 AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3291 GATGGTATGTTGATGTAGTGTAGTGAAGATGACCCGAAC 3329
RESULT 12
AAD49001
ID AAD49001 standard; cDNA; 3447 BP.
XX AAD49001;
XX
XX 07-MAR-2003 (first entry)
XX
XX Mouse MATER cDNA.
XX
XX Mouse; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.
XX
XX Mus sp.
XX
XX Key
XX CDS
XX Location/Qualifiers
9..3344
/*tag= a
/product= "Mouse MATER protein"
/transl_except= (pos:648..650, aa:Thr)
/transl_except= (pos:714..716, aa:Thr)
/transl_except= (pos:768..770, aa:Thr)
/transl_except= (pos:882..884, aa:Thr)
/transl_except= (pos:1359..1361, aa:Thr)
/transl_except= (pos:1554..1556, aa:Thr)
/transl_except= (pos:1740..1742, aa:Thr)
/transl_except= (pos:1869..1871, aa:Thr)
/transl_except= (pos:2046..2048, aa:Thr)
/transl_except= (pos:2049..2051, aa:Thr)
/transl_except= (pos:2229..2231, aa:Thr)
/transl_except= (pos:3216..3218, aa:Thr)
/transl_except= (pos:3300..3302, aa:Thr)
/transl_except= (pos:3330..3332, aa:Thr)
/transl_except= (pos:3333..3335, aa:Thr)
/notes= "No start and stop codon"
/partial
WO200281492-A1.
17-OCT-2002.
```

```
PF 29-MAR-2002; 2002WO-US009776.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2003-058494/05.
XX
XX P-PSDB; AAE31748.
XX
XX New isolated variant MATER proteins and nucleic acids, useful for
XX diagnosing, prognosing and treating infertility and reduced fertility,
XX and as contraceptive agents.
XX
XX Example 22; Page 83-85; 110pp; English.
XX
XX The present invention relates to novel MATER proteins and polynucleotides
XX encoding such proteins. The MATER proteins are essential to fertility.
XX Sequences of the invention are useful for diagnosing, prognosing and
XX treating infertility, reduced fertility and as contraceptive agents. They
XX are also useful in gene therapy. The method is useful for detecting a
XX predisposition to or pre-symptomatic screening of an individual for
XX infertility or reduced fertility. The present sequence is mouse MATER
XX cDNA
XX
XX Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,94e-204 Length: 3447
Score: 2688.00 Matches: 554
Percent Similarity: 66.49% Conservative: 186
Best Local Similarity: 49.78% Mismatches: 293
Query Match: 38.00% Indels: 81
DB: 7 Gaps: 11
US-10-066-521-6 (1-1344) x AAD49001 (1-3447)
QY 15 GlnTrpCysLeuTrpGluLeuAspLysGluGluPheGlnThrPheLysGluLeuLys 34
Db 163 CAATGGGTCTCCA-GAAAAGACAGTAAGCAATCTTGAAGCACCGTGGATTGGAAGAG 221
QY 35 LysLysSerSerGluSerThrCysSerSileProGlnPheGluIleGluAsnAlaAsn 54
Db 222 GAACAGAGAGTCAGAAAGACAC----- 242
QY 55 ValGluCysLeuAlaLeuLeuHisGluTyTrpGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
QY 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
Db 243 ---ATGCTCTCTTCAGAAATGTCTAGTAGAGCAATCTCTGAAGACAGCTGGATCAGAAGAA 299
QY 95 MetLysLysLysSerGln----- 100
Db 300 GTGGAAACAGCGTCAGAAAGAAAAATGACTTCTCCAGAAAAACGACGATAAATCAATCCAG 359
QY 101 -----AlaMetGluGlnGluGluAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAAGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 419
QY 116 GluIleSerGlnAlaMetGluGlnGluGluAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTCAGACAGCA-----GATAAGAT 443
QY 136 HisGlyGlyAspThrTrpAspTrpLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 444 AATGGAGGTGACTTACAAGACTACAGGCCCATGTGATTGCTAAGTTTCGACACACAGGTGTG 503
QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCCGCCAGAGATGAAATTTTGTCT 542
```



```
Db 2691 CGCGTGATCACTGAATCACTGCAACATTTAGATGATGCTTATGGCTTCTCGCAATGAGA 2750
Qy 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuMetAsnProValGluAspAsn 931
Db 2751 CTGCAACACACAAAGCTGACCCACTGAGCCTGACCATGACATGACCCCGTAGGGGATGGT 2810
Qy 932 GlyValLysLeuLeuCyGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGCTACTGTGTGAAGCTTTAAAGGAACCTACTTGTACCTTCAAGAACTGGAA 2870
Qy 952 LeuValLysCysHisLeuThrAlaAlaCysGlySerGluSerLeuSerCysValLysSerArg 971
Db 2871 CTAGTGGAGCTGCAACTCACACAGAACTGCTCGAGGACCTGGCCTGTATGATCACACA 2930
Qy 972 SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAAGCACTTAAAGTTTGGATCTTGTAAACAGCCCTGGGTGACAAAGGAGTCAATA 2990
Qy 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGGACTGAAGCAAAAGTAGCAGCTCCCTGAGGAGACTTGGGTGGGGCA 3050
Qy 1012 CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAAGTTGACTTCCAATTGCTGTGAGGCATTGTCTTGGCCATCTCTTGCACCCCTCAC 3110
Qy 1032 LeuThrSerLeuAsnLeuValGlnLysGlnLysAsnSerProLysGlyMetLysLeuCys 1051
Db 3111 CTGAACAGCTTAAACCTGTGGAAGTAATGACTTCAGTACATCGGGATGTTGAAGCTGTGC 3170
Qy 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyr 1071
Db 3171 TCTGCTTCCCAATGCCCTGTCTCTAACTGGGGATAATTGGCCTGTGGAAGCAGGAGTAC 3230
Qy 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIle 1091
Db 3231 TATGCCCGAGTGAGGAAGACAGCTGAGGAAGTTGAGTTTGTCAAGCCCAAGCTGTGTGATT 3290
Qy 1092 AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3291 GATGTGATTGGTATGCTAGTGTGAAGATGACCCGAAAC 3329
```

RESULT 13

```
ABK48609
ID ABK48609 standard; cDNA; 1157 BP.
XX AC ABK48609;
XX DT 13-AUG-2002 (first entry)
XX DE Human MATER cDNA fragment #1.
XX KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW Maternal antigen that embryos require; WATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX OS Homo sapiens.
XX FH Key
XX CDS 1. .1155
FT FT Location/Qualifiers
FT FT /tag= a
FT FT /product= "Human MATER"
FT FT /partial
FT FT /note= "No start or stop codon shown"
XX XX WO200232955-A1.
XX PN 25-APR-2002.
XX XX 04-APR-2001; 2001WO-US010981.
XX PF
XX XX
```

```
PR 18-OCT-2000; 2000US-0241510P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nelson LM, Tong Z;
XX DR WPI; 2002-454595/48.
XX PT P-FSDB; AAU79523.
XX FT
XX PT New isolated human Maternal Antigen That Embryos Require protein and
XX PT polynucleotide, useful in diagnosing or treating fertility or reduced
XX PT fertility, or as a contraceptive.
XX PT Claim 11; Page 68-69; 93pp; English.
XX CC The invention discloses an isolated human MATER (Maternal Antigen That
XX CC Embryos Require) protein, which can complement a WATER null phenotype in
XX CC which zygotes arising from the oocyte do not progress beyond the two-cell
XX CC stage. MATER is required for early embryonic survival and abnormal levels
XX CC of the protein can lead to premature ovarian failure (POF) and can be
XX CC caused by under or over expression of MATER or an autoimmune response to
XX CC MATER. MATER is a single-copy Maternal effect gene found on chromosome
XX CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
XX CC diagnosing or treating fertility and reduced fertility. In particular,
XX CC the MATER protein is useful as a contraceptive agent, or for influencing
XX CC (either inhibiting or enhancing) fertility and can be used to detect a
XX CC predisposition to infertility or reduced fertility, or for presymptomatic
XX CC screening of an individual for infertility/reduced fertility. The protein
XX CC and polynucleotide are also useful for detecting an excess or deficiency,
XX CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
XX CC human or a mouse) or for screening for a compound useful in influencing
XX CC MATER-mediated fertility. The sequence presented is the human MATER cDNA
XX CC fragment #1, which was isolated from a human cDNA library
SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.: 2.16e-150 Length: 1157
Score: 2005.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.34% Indels: 0
DB: Gaps: 0
```

US-10-066-521-6 (1-1344) x ABK48609 (1-1157)

```
Qy 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGly 271
Db 1 CAAGCTCCGGTGACGGAGATCATGTCCGACCAAGAGGCTGTGTTCATCATTTGACGGT 60
Qy 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
Db 61 TTCGATGACCTGGGCTCTGTCTTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG 120
Qy 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
Db 121 AAGCAGCCTCGTTTCACTCATACCGCAGTCTGTCTGAGGAAGGTCCTGCTCCCTGAGTCC 180
Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTGTCT 240
Qy 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu 351
Db 241 CCCGTTACCTGTAGTTAGAGGAATCTCCGGGACCAAGAAATCCACTTGTCTCTTGAG 300
Qy 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
Db 301 GCGGGATTTGGTGAGCATCAGACACAAAGGTTGCGTGCATCATGAAACCGTGAG 360
Qy 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACAGTCCAGGTGCCCGCGGTGCTCTCTCATCTCGGTGGCCCTTCAGCTG 420
```



```
QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
DB 421 CAGGACGTGGTGGGAGAGCGTCCGCCCTTCAACCAACGCTCACAGGCTCGACGCC 480
QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValAlaArgCysLeuAsnLeuGlu 431
DB 481 GCTTTTGTTCATCAGCTCACCCCTCGAGCGGTGGTCCGGCGCTGTCTCAATCTGGAG 540
QY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
DB 541 GAAAGAGTTGTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600
QY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyLeuSerGluLeuArgAla 471
DB 601 TCAGTGTGTGAGCGGTGAGCAGCTCATGTTTCAAGACTTCGGGGAGTCTGAGCTCCGTGCT 660
QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyTrpThrPhePhe 491
DB 661 CTGTTTTCATGAAATCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTC 720
QY 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyTrpValLeuGluGlyLeuGluIle 511
DB 721 CACCTCAGTCTCCAGGACTTCTGTCCGCCCTTGTACTACGTGTAGAGGCGCTGGAATC 780
QY 512 GluProAlaLeuCysProLeuTyTrpValGluLysThrLysArgSerMetGluLeuLysGln 531
DB 781 GAGCCAGCTCTCTGCCCTCTGTAGTGTGAGAAGCAAGAGGTCCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
DB 841 GCAGGCTTCCATATCCACTCGCTTGGATGAAGCGTTTCTGTGTTGGCTCGTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
DB 901 GACGTAAAGAGGCCACTGGAGTCTCTGGGCTGTCCCGTTCCTCCCTGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
DB 961 AAGCTTCTGCACCTGGGCTCTCTGTGGTTCAGAGCTTAATGCCACCCAGGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
DB 1021 ACCCTGGAGCGCTTCCACTGCTTTTCGAGACTCAGACCAAGAGTTTGTTCGCTTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
DB 1081 TTAACACAGCTTCCAAGAAGTGTGGCTTCCGATTAAACCAGAACCTGGACTTGTATGCATCT 1140
QY 632 SerPheCysLeuGln 636
DB 1141 TCCTTCTGCCCTCCAG 1155
RESULT 14
AAD48999
ID AAD48999 standard; cDNA; 1157 BP.
XX
AC AAD48999;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human MATER cDNA fragment 1.
XX
KW Human; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1155
FT /tag= a
FT /product= "Human MATER protein fragment 1"
FT /transl_except= (pos:112..114, aa:Thr)
```

```
FT /transl_except= (pos:589..591, aa:Thr)
FT /transl_except= (pos:865..867, aa:Thr)
FT /transl_except= (pos:973..975, aa:Thr)
FT /transl_except= (pos:1102..1104, aa:Thr)
FT /note= "No start and stop codon"
XX
XX /partial
PN WO200281492-A1.
```

XX

XX 17-OCT-2002.

XX 29-MAR-2002; 2002WO-US009776.

XX 04-APR-2001; 2001WO-US010981.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nelson LM, Tong Z;

XX WPI; 2003-058494/05.

XX P-PSDB; AAE31746.

XX New isolated variant MATER proteins and nucleic acids, useful for diagnosing, prognosing and treating infertility and reduced fertility, and as contraceptive agents.

XX Example 2; Page 79-80; 110pp; English.

XX The present invention relates to novel MATER proteins and polynucleotides encoding such proteins. The MATER proteins are essential to fertility. Sequences of the invention are useful for diagnosing, prognosing and treating infertility, reduced fertility and as contraceptive agents. They are also useful in gene therapy. The method is useful for detecting a predisposition to or pre-symptomatic screening of an individual for infertility or reduced fertility. The present sequence is human MATER cDNA fragment 1

XX SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,16e-150	Length:	1157
Score:	2005.00	Matches:	385
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	28.34%	Indels:	0
DB:	7	Gaps:	0

US-10-066-521-6 (1-1344) x AAD48999 (1-1157)

QY 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuPheIleIleAspGly 271

DB 1 CAAGCTCCGGTGCAGAGATCATGTCCTCCAGCAAGAGCTGTGTTCATCATTCACGCT 60

QY 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291

DB 61 TTCGATGACCTGGCTCTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG 120

QY 292 LysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311

DB 121 AAGAGCCTCCGTTCAACCTCATACGAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCC 180

QY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331

DB 181 TTCCTGATCGTCCCGTCAGAGACGCTGGGCACAGAGAACTCAAGTCAGAGGTCGTGCT 240

QY 332 ProArgTyTrpLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu 351

DB 241 CCCCCGTTACCTGTAGTAGGAATCTCCGGGGCAACAAAGAAATCCACTTGTCTCTTGAG 300

QY 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371

DB 301 CGCGGGATTGGTGAGCATCAGAAAGACACAGAGGTTCCGTCGCCATCATGAACACCGTGAG 360


```
QY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACGAGTGCAGGTGCGCGCTCTCTCATCTGCGTGGCCCTGCAGCTG 420
QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
Db 421 CAGGACGTGGTGGGGAGAGGTGCGCCCTTCAACCAAGCGCTACAGGGCTGCAGGCC 480
QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgCysLeuAsnLeuGlu 431
Db 481 GCTTTTGTGTTTCATCAGCTCACCCCTCAGGCGGTGTCGGCGCTGCTCAATCTGGAG 540
QY 432 GluArgValValLeuLeuArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
Db 541 GAAAGAGTGTCTCGAAGCGCTTCTGCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600
QY 452 SerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
Db 601 TCAGTGTTTGACGGTGACGACTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCT 660
QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePhe 491
Db 661 CTGTTTTCATGACATCTCTTCCACAGACCACTGTGAGGAGTACTACACTTCTTC 720
QY 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIle 511
Db 721 CACCTCAGTCTCCAGGACTTCTGTGCGCTTGTACTACGTGTAGAGGGCTGGAAATC 780
QY 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAGACAAAGAGGTCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCCATCTCGTTGGATGAGCGTTTCTTGTGGCTCTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GACGTAAAGAGGCCACTGGAGTCTCTGTGGCTGTCCCGTTCCTCGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGCATGGTCTCTCTGTGGTGCAGAGCCCTAATGCCACCACCCAGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGGAGCGCTTCCACTGTCTTTCGAGACTCAAGACAAAGAGTTGTTCGTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAAACAGCTTCCAAGAGTGTGGCTTCGATTAACCAAGAACCTGGACTTCATAGCATCT 1140
QY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 15
AAL4366
ID AAL44366 standard; cDNA; 3368 BP.
AC AAL44366;
XX
DT 31-OCT-2002 (first entry)
XX
DE Human PYRIN-3 cDNA sequence.
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-3.
```

```
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 70..3054
FT /*tag= a
FT /product= "Human PYRIN-3 protein"
XX
XX WO200261049-A2.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2002; 2002WO-US002967.
XX
XX 31-JAN-2001; 2001US-0265231P.
XX
XX 10-SEP-2001; 2001US-0318645P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX (AMHP ) WYETH.
XX
XX Bertin J, Wang W, Blatcher M;
XX
XX WPI; 2002-627477/67.
XX P-PSDB; AAO15593.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
XX diagnosing stress-related, apoptotic and inflammatory responses, or for
XX treating inflammatory and immune system disorders, cancers, or
XX neurological diseases.
XX
XX Claim 4; Fig 11; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
XX useful for modulating and diagnosing stress-related, apoptotic and
XX inflammatory responses. The PYRIN protein and DNA sequences are useful
XX for treating: inflammatory disorders and immune system disorders (e.g.
XX Crohn's disease, psoriasis, graft rejection, multiple sclerosis, contact
XX dermatitis, reactive arthritis, allergies, viral infections and
XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
XX (e.g. systemic lupus erythematosus and arthritis); and neurological
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
XX protein and DNA sequences may also be used in screening assays, detection
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
XX predictive medicine (e.g. diagnostic assays, clinical trials and
XX pharmacogenomics) and transcription profiling. The present DNA sequence
XX encodes the human PYRIN-3 protein
XX
XX SQ Sequence 3368 BP; 900 A; 803 C; 828 G; 837 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.62e-104 Length: 3368
Score: 1435.00 Matches: 349
Percent Similarity: 48.05% Conservative: 204
Best Local Similarity: 30.32% Mismatches: 448
Query Match: 20.29% Indels: 150
DB: 6 Gaps: 15

US-10-066-521-6 (1-1344) x AAL44366 (1-3368)
QY 9 PheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThr 28
Db 85 TTCTCTGATTTGGTCTTATGTGGTATCTGGAGGAGCTCAAAAAGGAGGAGTTCAGGAA 144
QY 29 PheLysGluLeuLeuLysLysSerSerGluSerThrThrCysSerIleProGlnPhe 48
Db 145 TTTAAAGACATCTCAAGCAAAATGACTTTCAGCTTCAAGCAGATTCCTCGACT 204
QY 49 GluIleGluAsnAlaAsnValGluCysLeuAlaLeuLeuHisGluTyrTyrGlyAla 68
Db 205 GAGGTCAAAAAGGATCCCGGGAAGAACTTGCAAACCTTTGTATCAAGCACTATGAAGAA 264
QY 69 SerLeuAlaTrpAlaThrSerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSer 88
```

Db	265	CAACAAGCTTGGAAACATAACTTTAAGAACTTTCAAAAAGATG-----	306
Qy	89	GIuLySaLaArGrAspAspMetLySLeSLeSerGlnAlaMetGluGlnGluGlyAlaThr	108
Db	307	-----GATAGAAAGGATCTCTGCATGAAGGTTCATGAGGAGAGACAGCA	351
Qy	109	AlaAlaGluThrGluGluGlnGluLeuSerGlnAlaMetGluGlnGluGlyAlaThrAla	128
Db	352	TACAACAAGACTTATCAAGCTCAGCAAAAGCAGAAATTCAGCCCG-----	396
Qy	129	AlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLySLeSerHisValMet	148
Db	397	-----TTATGGTCCAGCAAGTCTGTCACTGAG	423
Qy	149	ThrLySLeAlaGluGluGluAspValArgArgSerPheGluAsnThrAlaAlaAspTrp	168
Db	424	ATTCACTATATCTTGGAGGAGAGTCAAGCAAGAA-----	459
Qy	169	ProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArg	188
Db	460	---GAATGTGACCATTTGGACCGCTTTTCTCCAAAGGAAACTGGGAAACAGCCACGT	516
Qy	189	ThrValValLeuHisGlyLySLeSerGlyLeGlyLySLeSerAlaLeuAlaArgAlaVal	208
Db	517	ACAGTGATATTCAAGGACCAACAAGGAATTTGGAAAAACGACACTCTGATGAAGCTGATG	576
Qy	209	LeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuPro	228
Db	577	ATGCGCTGGTCGGACAACAAGATCTTTCCGGATAGTTCCTGTACACGTTCTATTCTGCG	636
Qy	229	ValArgGluMetGlnArgLySLeGluSerSerValThrGluPheLeuSerArgGluTrp	248
Db	637	TGCAGAGAACTGAGGGAGTTGCCGCCAACAGAGTTTGGCTGACTTGAATTTCCAGAGAGTGG	696
Qy	249	ProAspSerGlnAlaProValThrGluLeuMetSerArgProGluArgLeuPheLeu	268
Db	697	CCTCACCCCGCTGCTCTATACAGAGATCGTGTCTCAACCGGAGAGACTCTCTGTCTGCTC	756
Qy	269	IleAspGlyPheAspAspLeuGlySerValLeuAsnAsn---AspThrLySLeuCysLyS	287
Db	757	ATCGACAGCTTCGAAGAGCTGCAGGGCGGTGTGAACCGCAATTCGGATCTGTGTGGT	816
Qy	288	AspTrpAlaGluLySLeGlnProProPheThrLeuLeuArgSerLeuLeuArgLySValLeu	307
Db	817	GACTTGATGAGAAACGGCGGTGCAGGTGCTTCTGAGCAGTTTGTCTGAGGAAGAAGATG	876
Qy	308	LeuProGluSerPheLeuLeuValThrValArgAspValGlyThrGluLySLeuLySLeuSer	327
Db	877	CTCCCGGAGGCTCCCTGCTCATCGCATCAAAACCGGTGTGCCGAAGGAGCTCCGGGAT	936
Qy	328	GluValValSerProArgTyrLeuLeuValArgGlyLeuSerGlyGluGlnArgLeuHis	347
Db	937	CAGGTGAGATCTCAGAAATCTACGCCCGGGGATTCACAGAGATGATAGTTAGTTAGT	996
Qy	348	LeuLeuLeuGluArgGlyIleGlyGluHisGlnLySLeuThrGlnGlyLeuArgAlaIleMet	367
Db	997	TATTTCTGCTGTTCTTCAAGACCCCGAAAGACCCATGGAGCGCTTCAATCTTGTAAGA	1056
Qy	368	AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLeuCysVal	387
Db	1057	GAAGTGAACAGCTGTTTTCATATGCAAAATCCCGCTCTCTGCTGATCCTGTGTACC	1116
Qy	388	AlaLeuGlnLeuGlnAspValValGlyLeuSerValAlaProPheAsnGlnThrLeuThr	407
Db	1117	AGTCTGAAGCAAGAGATGCGAAAGGAAAGACCTGGCCCTTGACCTGCCAGAGACTATCC	1176
Qy	408	GlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCys	427
Db	1177	TCGTGTACTCTCTTCTGCTTTTAACTGTTTCACACTGAGGGTCCGAGGGCCGACT	1236
Qy	428	LeuAsnLeuGluGluArgValValLeuLySLeuArgPheCysArgMetAlaValGluGlyVal	447

Search completed: July 16, 2004, 00:59:44
Job time : 1282 secs

```
QY 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2233 -----GCAGCAAC----- 2241
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2242 -----GTCAAAGAGCTAGCGCTGGTA 2262
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2263 AATTGTCACCTCTACCCATTGATTGTAAGTCTTGCTGGCCTTTAACCAACAACAAG 2322
QY 860 SerLeuThrHisLeuLysSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2323 AAGCTGACGTATCTGAATGATCTCTGCAACCACTTA---GACACAGGCGTGCCTTTTG 2379
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2380 TGTGAAGCCCTGTGCAGCCCAACACACACGCTCTGGTATACCTGTATGTTGGCTTTCTGCCAC 2439
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2440 CTCAGCGAGCAGTCTGCAATACATCTCTGAATGCTTCTGCGTAAACAGAGCGTGGC 2499
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLysGluVal 939
Db 2500 TATCTAGACCTCAGTGCCTCAATGCTCTGAAGGACGAGGAGCTGAAACTCTCTGCGAGGCC 2559
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 2560 TTGAACATCCGGACTGCTGCTGGATTCACTGTGTTGGTAAATGTTTATCACTGCT 2619
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 2620 GCTGCTGTGAAGACCTCGCTCTGCTCATCAGCAATCAAAACCTGAAGATTCTGCAA 2679
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaLeuLysGluGlyValLysGln 999
Db 2680 ATTGGTGCATGAATCGGAGATGTGGTGTGCAGCTGTGTGTCGGGCTCTGACGCAT 2739
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 2740 ACGGATTCCCGCTAGAGATTCTTGGGTGGAAGAAATGCGGTTAAACGAGCACCTGCTGT 2799
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 2800 AAGGATCTCGCGTCTGTTCTCACCTGCAGTAAGACCCCTGCAGCAGCTCAACCTGACCTTG 2859
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuLysSerAlaPheAlaCysProThrSer 1059
Db 2860 AACACCTTGGACACACAGGGGTGGTGTACTCTGTGAGGCGCTGAGACACCCAGAGTGT 2919
QY 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 2920 GCCCTGCAGGTCTCGGCTCGAGAAACTGATTTGATGAGAAACCCAGGCACCTTCTG 2979
QY 1080 GluGluValGlnLeuLysProArgValIleLeuAsp----- 1092
Db 2980 ACGGCTGAGGAAGAGAGAAATCCTTAACCTGACCATCAGATGACTGTGACACCAATCACA 3039
QY 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLys 1105
Db 3040 AGGGTAGAGATCTGATTTCGGAGGAACCTGG----- 3069
QY 1106 IleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGly 1125
Db 3070 ---GCTCTGACTCGAAACACCTGCAAGGACAGGAGCTGGAC---CGTTACTTACATGAC 3123
QY 1126 MetAsnProGluGlnLysLysArgValSerLeu 1136
Db 3124 ACTGCACCCGAGGATACAAATCATTTGACACTC 3156
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 20:38:28 ; Search time 200 Seconds
(without alignments)
3729.270 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGDKSLTFSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spoel/US10066521/runat 13072004 122213 11111/app query.fasta_1.1543
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -STARTP=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 @CGN 1 1 107 @runat 13072004 122213 11111 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=1120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699.5	9.9	1566	US-09-833-381-1975	Sequence 1975, Ap
2	630.5	8.9	1374	US-08-910-731-3	Sequence 3, Appli
3	630.5	8.9	1374	US-08-795-395-3	Sequence 3, Appli
4	621	8.8	1371	US-08-910-731-1	Sequence 1, Appli
5	621	8.8	1371	US-08-795-395-1	Sequence 1, Appli
6	612	8.7	1371	US-08-910-731-7	Sequence 7, Appli
7	607.5	8.6	1386	US-08-910-731-5	Sequence 5, Appli
8	458	6.5	1515	US-09-833-381-1904	Sequence 1904, Ap
9	458	6.5	1515	US-09-833-381-1505	Sequence 1905, Ap
10	394.5	5.6	1650	US-09-833-381-1976	Sequence 1976, Ap
11	379.5	5.4	4141	US-09-245-281-42	Sequence 42, Appl
12	379.5	5.4	4141	US-09-207-359B-42	Sequence 42, Appl

13	379.5	5.4	4141	4	US-09-340-620A-42	Sequence 42, Appl
14	379.5	5.4	4141	4	US-09-865-364-42	Sequence 42, Appl
15	359.5	5.1	3382	4	US-09-099-041A-7	Sequence 7, Appli
16	359.5	5.1	3382	4	US-09-245-281-7	Sequence 7, Appli
17	359.5	5.1	3382	4	US-09-207-359B-7	Sequence 7, Appli
18	359.5	5.1	3382	4	US-09-340-620A-7	Sequence 7, Appli
19	359.5	5.1	3382	4	US-09-865-364-7	Sequence 7, Appli
20	347	4.9	2859	4	US-09-099-041A-9	Sequence 9, Appli
21	347	4.9	2859	4	US-09-245-281-9	Sequence 9, Appli
22	347	4.9	2859	4	US-09-207-359B-9	Sequence 9, Appli
23	347	4.9	2859	4	US-09-340-620A-9	Sequence 9, Appli
24	347	4.9	2859	4	US-09-865-364-9	Sequence 9, Appli
25	327	4.6	4543	2	US-08-519-547A-5	Sequence 5, Appli
26	319.5	4.5	4366	4	US-09-064-199-12	Sequence 12, Appl
27	319.5	4.5	4366	4	US-09-064-199-14	Sequence 14, Appl
28	319.5	4.5	4418	4	US-09-064-199-13	Sequence 13, Appl
29	319.5	4.5	4431	4	US-09-064-199-8	Sequence 8, Appli
30	319.5	4.5	4441	4	US-09-064-199-2	Sequence 2, Appli
31	319.5	4.5	4441	4	US-09-064-199-10	Sequence 10, Appl
32	319.5	4.5	4549	4	US-09-064-199-9	Sequence 9, Appli
33	319.5	4.5	4564	4	US-09-064-199-2	Sequence 2, Appli
34	319.5	4.5	4649	4	US-09-064-199-11	Sequence 11, Appl
35	319.5	4.5	4746	4	US-09-064-199-7	Sequence 7, Appli
36	319.5	4.5	5105	4	US-09-064-199-3	Sequence 3, Appli
37	319.5	4.5	5463	4	US-09-064-199-1	Sequence 1, Appli
38	309	4.4	3393	1	US-08-295-502-1	Sequence 1, Appli
39	309	4.4	3393	5	PCT-US95-10691-1	Sequence 1, Appli
40	307	4.3	4302	4	US-09-245-281-38	Sequence 38, Appl
41	307	4.3	4302	4	US-09-207-359B-38	Sequence 38, Appl
42	307	4.3	4302	4	US-09-340-620A-38	Sequence 38, Appl
43	307	4.3	4302	4	US-09-865-364-38	Sequence 38, Appl
44	296.5	4.2	2480	4	US-09-064-199-15	Sequence 15, Appl
45	235	3.3	3116	4	US-09-904-615-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-833-381-1975
; Sequence 1975, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1975
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1975

Alignment Scores:
Pred. No.: 1.51e-63 Length: 1566
Score: 699.50 Matches: 178
Percent Similarity: 51.56% Conservative: 103
Best Local Similarity: 32.66% Mismatches: 201
Query Match: 9.89% Indels: 63
DB: 4 Gaps: 15

US-10-066-521-6 (1-1344) x US-09-833-381-1975 (1-1566)

Oy	139	AspThrTrpAspTyr-----LysSerHisValMetThrLysPheAlaGluGlu 154
Db	19	GAGACATGGGACACATCAGTTCGCTTAAGACCATATATATCCGT----- 66
Oy	155	GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrIeu 174

67 -----AATACATCAAGGAC-----GAACATGAGGAACGTG 96
175 AlaGlyAlaPheAspSerAspArgTyrGlyPheArgProArgThrValValLeuHisGly 194
197 CAGGCGCTTACTGATCCTAATAGGACTAGAGCCAGGCCAGACGATAGTCTTGGTGGG 156
195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTyrAlaGlnGly 214
157 AGGCGAGGGTGGAGAGACCACCTTGGCAATGAGGCTATGTCACCTGGGCAATAGGA 216
215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
217 GTTCTCTTTCAGCAAGGTTCTCTATGTTTTCTATCTCAGCTCCCATATAAAGGTAC 276
235 LysLysGluSerValThrGluPheIleSerArgGluTyrProAspSerGlnAlaPro 254
277 ATGAAGAAATACACTTGTGCTGAATGATTTCTTGTGATTGGCCGATTTGTATGCC 336
255 ValThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAsp 274
337 ATTGAAGAGTTGATGCTCAACACAGAGAGCTCTGTATTATTTGATGGCTTTGAGAA 396
275 Leu-----GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTyr 289
397 ATAATCATATCTGAGTCAGCTCTGAGAGCTTGGATGATGCTCGCCATGTACAGACTGG 456
290 AlaGluLysGlnProProPheThrLeuIleArgSerLeuArgLysValLeuLeuPro 309
457 TACCAGGAGCTCCAGTGCACCAATCTACACAGCTTGTGTGAAGAAGAAATTTGGTCCC 516
310 GluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluVal 329
517 CTGGCTACCTTACTGATGATCAGACCTGCTGTTGTGAGATCTTAAGGCGCTCAATTA 576
330 ValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGln---ArgIleHisLeu 348
577 GTGAATCCATGTTTGTACAAATTAACAGGTTTACAGGGGACGACCTACGGGTATATTC 636
349 LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGly---LeuArgAlaIleMet 367
637 ATG-----AGACACTTGTGATGACTCAAGTGAAGTTGAGAAAATCTTCGACGAGCTAAGA 690
368 AsnAsnArgGluLeuLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal 387
691 AAAAAGCAAACTCTCTTTCATCTGTCAGTCCCGCCCATGTTGTGGACCGTATGTTCC 750
388 AlaLeuGlnLeuGlnAspValValGlyLysValAlaProPheAsnGlnThrLeuThr 407
751 TGTCTGAAGCAGCCGAGGAGGTGAGGTATTACGATCTCCAGTCAATCACTCAGACTACCACC 810
408 GlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArg----- 425
811 AGTCTGTATGCTTATTTT---TTCTCCAACCTTGTCTCCACGACGAGGTAGATTTGGCA 867
426 -----ArgCysLeuAsnLeuLeuGluArgValValLeuLys 437
868 GATGACAGCTGGCCAGACAAATGAGGCGCCCTC----- 900
438 ArgPheCysArgMetAlaValGluGlyValTyrAsnArgLysSerValPheAspGlyAsp 457
901 -----TGCAGTCTGGCCATAGAAGGCTGTGGTCTATGAACCTTCACTTTAACAAGAA 954
458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
955 GACACTGAGATCGAGGCGCTGGAAGTGCCTTTTCAATGATTCTCTCAGAGTTCAATATT 1014
478 LeuLeuProAspSerHisCysGluGluTyrTyrPhePheHisLeuSerLeuGlnAsp 497
1015 CTTCAAAGATCAATGACTGGGGGTTGCACTACTTTTCCACCCACCTAAGTTCCAGAG 1074
498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluLeuGluProAlaLeuCysPro 517
1075 TTTTTCGAGCCATGCTCTTGTGCTAGAG-----GAACTAGAGAAATTCCT 1122

Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIle--- 536
Db 1123 CCCCATTCACAAAGCACAA-----GAGATGAAGATGTTACTGCAACACGCTCTTG 1173
Qy 537 -----HisSerLeuTyr-----MetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 1174 CTTGACAAAGACCTTACGACTCCAGCTGCTCTTCTTCTTGGTCTTTTAAATAAA 1233
Qy 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 1234 AACATGACGAAGAACTGGAGATACTTTTGCAATGTAAATAATCTCCAGGGTAATGGAG 1293
Qy 572 LysLeuLeuHisTyrValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 1294 GAATTATTAAGTGGGAGAGAGTTAGGTAAAGTGAAGTGCCTCTCTCCAATTTCCAC 1353
Qy 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1354 ATTCTACGACTTTTTCACCTGCTTACAGAGTCCAGAGGAGAGACTTTCACAAAGAGATG 1413
Qy 612 LeuAsnSerPheGlnGluValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1414 TTGGGTGCTATCTTTGAAGTTGACCTTAATAATTTTGGAGGACGAAGAACTCCAAGCTTCT 1473
Qy 632 SerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGly--- 650
Db 1474 TCATTTTGGCTAAAGCACTGTAAAGGTTAAATAAGCTTAAGGCTTTCTGTGTAGCAGTCCAC 1533
Qy 651 IlePheProArgAsp 655
Db 1534 ATCTTGAAGGGAC 1548
RESULT 2
US-08-910-731-3
; Sequence 3, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 16-AUG-1996
; APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1374 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1368
 US-08-910-731-3

Alignment Scores:

Pred. No.: 2,41e-56 Length: 1374
 Score: 630.50 Matches: 151
 Percent Similarity: 51.38% Conservative: 73
 Best Local Similarity: 34.63% Mismatches: 205
 Query Match: 8.91% Indels: 9
 DB: 2 Gaps: 3

US-10-066-521-6 (1-1344) x US-08-910-731-3 (1-1374)

```

Qy 653 ProArgpGluSerAlaGluAlaCysProValProLeuTrpMetArg-----Asp 670
Db 39 CCGTGGACAGAGCTCTTCCCTGATCCAAATACCAAG-TGGTC-AGGCTGGATGAC 96
Qy 671 LysThrLeuLeuGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHis 690
Db 97 TGTGGCTCCTCAAGTGGGTGGCAAGACATCAGGTGAGCATCCAGGCCAACCTGCC 156
Qy 691 LeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluAlaMetLeuThrLeuCys 710
Db 157 CTGACAGAGCTCAGGCTACGCACCAATGAATCGGTGATGCTGGTGGTCTGGTCTC 216
Qy 711 AlalysLeuArgHisProThrCysIleGlnThrLeuMetPheArgAsnAlaGlnIle 730
Db 217 CAGGCGCTGCAGAAATCCCACTTGTGAAGATCCAGAAGCTGAGCCCTTCAGAACTG 276
Qy 731 Thr---ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSer 749
Db 277 ACGGAGCTGGCTGGGGTCTGCTGATGCTGGCTCTTGTCTACCTCGGTGAA 336
Qy 750 LeuAsnLeuGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeu 769
Db 337 CTACATCTCAATGACAACTCTGGGGATGAAGCCCTGAAGCTGCTCTGTGAAGGACTC 396
Qy 770 LysHisProLysCysLeuLeuLysSerLeuArgLeuAspCysGlyLeuThrHisAla 789
Db 397 CGGGACCCCAAGTCCGCTTGTGAAGCTTCAAGTGAATCTGTAACCTCACAGCTACC 456
Qy 790 CysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeu 809
Db 457 AGTGGGAGCCCTGGCTGCTAGTCTAGGTGAAACCTGACTTAAAGACTAGTATG 516
Qy 810 AlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSer 829
Db 517 AGCAACAATGATCTTCATGAGCTGGTATCCACACTCTGTGCCAGGGCCCTGAAGATCT 576
Qy 830 GlnCysAlaLeuGlnLysLeuLeuLeuGluAspCysGlyIleThrAlaThrGlyCysGln 849
Db 577 GCCTGTCAACTGGAGTCACTCAAACTGGAGAACTGTGTATACATCAGCCAACTGCAAG 636
Qy 850 SerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsn 869
Db 637 GATCTGTGTGATGTTGGCTCCAAAGCCCTCACTGCAAGAACTGGACTTGGGAGCAAC 696
Qy 870 SerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSer 889
Db 697 AAGCTGGGCAACACAGGCACTGTGCTCAGGACTGCTCTCCAGCTGCAGG 756
Qy 890 LeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAla 909

```

```

Db 757 CTGAGGACTCTGTGGCTCTGGGACTGTGATGTCACTGCGAAGGCTGCAAGGACCTGTGC 816
Qy 910 LeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGlu 929
Db 817 CGTGTCTCAGAGCCAAAGCAGAGCCTGAAGAACTAGCCTAGCTGGCAATAGCTGAAG 876
Qy 930 AspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAsp 949
Db 877 GATGAGGCTGCCCAACTGCTGTGTAGAGCCTGTGTAGAGCCTGGCTGTACGTGAGTCA 936
Qy 950 LeuGluLeuValLysCysHisLeuThrAlaAlaCysGlySerLeuSerCysValIle 969
Db 937 CTGTGGGTAAAGACCTGTAGCCTCAGAGCTGCTCTTGTCCCACTTCTGCTCGGTGTG 996
Qy 970 SerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGly 989
Db 997 ACCAAACACAGTTCTCTGTTTGTGTTGCAATGAGCAGCAACCCGCTGGGAGACTCGGA 1056
Qy 990 ValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeu 1009
Db 1057 GTGTGGAGCTTTGCAAGGCCCTGGGCTATCCGACACAGTGTGCTGTGCTTGTGCTG 1116
Qy 1010 LysAlaCysGlyLeuThrSerAspCysGlyAlaLeuSerLeuAlaLeuSerCysAsn 1029
Db 1117 GAGACTGTGATGTGACAGACAGTGGCTGAGCAGCCTTCCACTGTCTGCTGGCCCAAC 1176
Qy 1030 ArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLys 1049
Db 1177 CGCAGCTTGAGGNACTGGACCTCAGTAACAACACTGCATGGGGGACACAGCGTGTCTACA 1236
Qy 1050 LeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu----- 1066
Db 1237 CTGTGAGAGCTCAAAACAGCCAGCTGCATCTTCCAGCAGCTTGTCTGTATGACATT 1296
Qy 1067 ---TrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1081
Db 1297 TACTGGACGATGAGTGGGAAGACCACTTCCGGCCCTGGAGGAGGAA 1344

RESULT 3
US-08-795-395-3
; Sequence 3, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600

```

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1374 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1368

US-08-795-395-3

Alignment Scores:

Pred. No.: 2,41e-56 Length: 1374

Score: 630.50 Matches: 151

Percent Similarity: 51.38% Conservative: 73

Best Local Similarity: 34.63% Mismatches: 205

Query Match: 8.91% Indels: 9

DB: 2 Gaps: 3

US-10-066-521-6 (1-1344) x US-08-795-395-3 (1-1374)

```

QY 653 ProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetArg-----Asp 670
DB 39 CGGGTGGACAGAGCTCTTCCCTGATCCACAAATACCAAG-TGGTC-AGGCTGGATGAC 96
QY 671 LysThrLeuLeuGluGluTrpGluAspPheCysSerMetLeuGlyThrHisProHis 690
DB 97 TGTGGCTCTCACTAGTGGTGGTCAAGACATCAGGTGAGGATCCAGGCAACCTGCC 156
QY 691 LeuArgGlnLeuAspLeuSerSerLeuLeuThrGluArgAlaMetLysThrLeuCys 710
DB 157 CTGACAGAGCTCAGCTACGACCAATGAACTGGGTGATGCTGGTGGTCTGGTCTC 216
QY 711 AlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIle 730
DB 217 CAGGCTGACAGATCCCACTTGAAGATCCAGAGCTGAGCCTTCAGACTGCGAGTTC 276
QY 731 Thr---ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSer 749
DB 277 ACGGAAGCTGGCTGGGGTCTCGCTGATGCTGGTGGTCTTTGTCTACCTGCGTGAA 336
QY 750 LeuAsnLeuGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeu 769
DB 337 CTACATCTCAATGATCAACCTCTGGGATGAGGCTGAGCTGCTGTGAAGGACTC 396
QY 770 LysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAla 789
DB 397 CGGAGCCCCAGTCCGCTCTTGAGAGCTTCACTTGAATACTGTAACTCAGACTACC 456
QY 790 CysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeu 809
DB 457 AGCTGCGAGCCCTGGCTGCTCAGGGTGAACCTGACTTTAAAGAGCTAGTATTG 516
QY 810 AlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSer 829
DB 517 AGCAACATGATCTCCATGAGCTGGTATCCACACTCTGTGCCAGGCCCTGAGGATCT 576
QY 830 GlnCysAlaLeuGlnLysLeuLeuGluAspCysGlyIleThrAlaThrGlyCysGln 849
DB 577 GCCTGTCACTGGAGTCACTCAAACTGGAGACTGTGTATCATCATCAGCCACTGCAAG 636
QY 850 SerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsn 869
DB 637 GATGTGTGATGTGTGGCTCCAAAGCCTCACTGCAAGAACTGGACTTGGGCGAGCAAC 696
QY 870 SerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSer 889
DB 697 AAGCTGGCAACACAGGCAATGACGACTGTGCTCAGGACTGTGCTCCAGCTGAGG 756
QY 890 LeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAla 909

```

```

DB 757 CTGAGGACTCTGTGGCTCTGGGACTGTGATGTCTCACTGCAGAGGCTGCAAGGACCTGTGC 816
QY 910 LeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGlu 929
DB 817 CGTGTCTCTCAGAGCCCAAGCAGAGCTTGAAGAACTCAGCCTAGCTGCAATGAGCTGAAG 876
QY 930 AspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAsp 949
DB 877 GATGAGGTGCCCAACTGTGTGAGAGCTTTAGAGCTCTTACGCTGGGATGCA 936
QY 950 LeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIle 969
DB 937 CTGTGGGTAAAGACCTGTAGCCTTCACAGCTGCCTTGTGCCCACTTCTGCTCGGTGTG 996
QY 970 SerArgSerArgHisLeuLysSerLeuAspLeuLeuThrAspAsnAlaLeuGlyGly 989
DB 997 ACCAAAACAGTCTCTCTGTTGAGTTGCAATATGACGAGCAACCCGCTGGGAGACTCGGA 1056
QY 990 ValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeu 1009
DB 1057 GTCTGGAGCTTTGCAAGGCCCTGGGCTATCCGGACACAGTGTGCTGTGCTTGGCTG 1116
QY 1010 LysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsn 1029
DB 1117 GGAGACTGTGATGTGACAGACAGTGGCTGCAGCAGCTTGGCACTGCTGCTGGGCCAAC 1176
QY 1030 ArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLys 1049
DB 1177 CGCAGCTTGGAGGAACTGGACCTCAGTAACACTGATGGGGGACCAACGCTGTCTTACAA 1236
QY 1050 LeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu----- 1066
DB 1237 CTCTGGAGAGCTCAAAACAGCCAGCTGATCTTTCAGAGCTTGTCTGTATGACATT 1296
QY 1067 ---TrpLysTrpGlnTrpValGlnIleArgLysLeuLeuGluGlu 1081
DB 1297 TACTGGACGATGAGTGGAGAGCAGCTTGGGGCCCTGGAGGAGGAA 1344

```

RESULT 4

```

US-08-910-731-1
; Sequence 1, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:

```


Db	607	AACTGGCGTCTCAAGCCAGCAAACTGTCGGAATTTGTGGCTCTCCAGGCC	666
Qy	860	SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu	879
Db	667	TCGCTGAGGAGCTTGACCTGGCAGCAACGGGCTGGGCGACGCGGCATAGCGGAGCTG	726
Qy	880	CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis	899
Db	727	TGCCCGGGCTCTTGAGCCCGGCTCCCGGCTCAAGACCCCTGTGGCTCTGGAGTGTGAC	786
Qy	900	LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr	919
Db	787	ATCACCGCCAGTGGCTGCAGAGACCTTGCCGCTGCTTCCAGGCCAAGAGACCTCAAG	846
Qy	920	HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuCysGluVal	939
Db	847	GAGCTCAGTCTGGCGGCAACAGCTGGGCGACGAGGGCGCCGGCTGTGTGCGAGAGC	906
Qy	940	MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla	959
Db	907	CTGCTGCAGCCCGGCTGCCAGCTGGAGTCCCTGTGGTGAAGTCTTCGACGCTCAGCGCG	966
Qy	960	AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp	979
Db	967	GCCTGTGTCAGCAGCTCAGCTTGATGCTGACCCAGAACAAAGCATCTCTGSACTTCAG	1026
Qy	980	LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln	999
Db	1027	TTGAGCAGCACAAAGCTGGTGACTTGGGACTCCAGGAGCTGTGCCAGGCCCTGAGCCAG	1086
Qy	1000	LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys	1019
Db	1087	CCGGCACCACACTGCGGGTGCTCTGCTTTGGGACTGTGAGGTGACCAACAGCGGCTGC	1146
Qy	1020	GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln	1039
Db	1147	AGCAGCTCGCCTCGCTGCTGCTGGCCAAACCGCAGCCTGCGAGAGCTGGACCTGAGCAAC	1206
Qy	1040	AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer	1059
Db	1207	AACTGTGGCGACCCGGCGCTCTGCGACTCTGGGGAGCTGGAGCAGCCGGGCTGC	1266
Qy	1060	AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyProValGlnIleArgLysLeuLeu	1079
Db	1267	GCCTTGAGCAGCTGGTCTGTACGACACACTACTGGCAGGAGGTGGAGGACCGGCTG	1326
Qy	1080	GluGluValGlnLeuLysPro-----ArgValVal 1090	
Db	1327	CAGGCGCTGGAGGGAGCAAGCCCGGCTGAGGGTCAATC 1365	

RESULT 5

US-08-795-395-1

; Sequence 1, Application US/08795395

; Patent No. 5965399

; GENERAL INFORMATION:

; APPLICANT: CHATTERJEE, DEB K.

; APPLICANT: SHANDILYA, HARINI

; TITLE OF INVENTION: Cloning and Expression of Rat Liver and

; NUMBER OF INVENTION: Porcine Liver Ribonuclease Inhibitor

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

; STREET: 1100 NEW YORK AVE., N.W., SUITE 600

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/795,395
 ; FILING DATE: 04-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/024,057
 ; FILING DATE: 16-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ESMOND, ROBERT W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0942.3440002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1371 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1368
 ; US-08-795-395-1

Alignment Scores:
 Pred. No.: 2,44e-55 Length: 1371
 Score: 621.00 Matches: 151
 Percent Similarity: 50.11% Conservative: 86
 Best Local Similarity: 31.92% Mismatches: 210
 Query Match: 8.78% Indels: 26
 DB: 2 Gaps: 5

US-10-066-521-6 (1-1344) x US-08-795-395-1 (1-1371)

Qy 625 AsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIle 644
 Db 4 AACCTCGATATT-----CATTCGAGCAGCTGAGGAGCGCC 39
 Qy 645 ArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValVal 664
 Db 40 CGG-----TGACAGAGCTCTCTCGCGCTGCTC 66
 Qy 665 ProLeuTrp-----MetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
 Db 67 CAGCAGTATGAGGTGGTCAGGCTCGACGCTCGCGGCTCGAGGAGGACACTGCAAGGAC 126
 Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerIle 700
 Db 127 ATCGGTCTGCGCTCCGGGCGCAACCTCTCTGACCGAGCTCTGCTCCGACCAACAGCAG 186
 Qy 701 LeuThrGluArgAlaMetLysThrLysCysAlaLysLeuArgHisProThrCysLysIle 720
 Db 187 CTGGCGATGCGCGGTGCACCTGGTGTGAGGCGCTCGAGGCGCTCGAGCCCTGCAAGATC 246
 Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGlyValGlnHisLeuTrpArg 739
 Db 247 CAGAAGCTCAGCCTCGAGAACTGTCTGACCGAGGCGGCTCGCGGGTCTCGCCCGC 306
 Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLysGlu 759
 Db 307 ACGTGGCTCCCTGCCCGCCAGCTCGGGAGTGTATCTCAGCGCAACACCCTGCGGGGAC 366
 Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
 Db 367 GCCGCGCTCGGCTGTCTGAGGGGCTCTCTGAGCCCGCCAGTGCACCTGGAGAGCTG 426
 Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysLysSerGlnIleLeuThr 799
 Db 427 CAGTTGAGTACTGCGCCCTGACCGCGCCAGCTGCGAGGCGCCCTGCGCTCGCTCAGG 486
 Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819

Db 487 GCCACGCGCGCTTGAGGAGCTCAGGTGAGCAACACGACATCGCGGCGCGCGCC 546
 Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
 Db 547 CGGTGTCTGGGCCAGGCTCTGGCAGACTCTGCTGCCAGCTGGAGAGCGCTCAGGCTGGAG 606
 Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 Db 607 AACTGCGGTCTCAGCCAGCAACTGTCGGAATTTGTGGCTCTCCAGGCC 666
 Qy 860 SerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
 Db 667 TCCTGAGGAGCTTGACCTGGCAGCAACGGGCTGGGCGACCGCGGCATACCGGAGCTG 726
 Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
 Db 727 TGCCCGGGCTCTTGAGCCCGGCTCCCGCTCAAGACCTGTGGCTCTGGGAGTGTGAC 786
 Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
 Db 787 ATCAGCGCCAGTGGCTGCAGAGACCTCTGCGGTGTCTCCAGGCCAAGGAGACCTGAAG 846
 Qy 920 HisLeuSerSerMetAsnProValGluAspAsnGlyValLysLeuCysGluVal 939
 Db 847 GAGCTCAGCTGCGGGGCAACAGCTGGGCGAGCGGCGCGCTGTGTGCGAGAGC 906
 Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
 Db 907 CTGCTCAGCGCGGCTGCCAGCTGGAGTCCCTGTGGGTGAAGTCTTGACGCTCAGCGCG 966
 Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
 Db 967 GCCTGCTGCCAGCAGCTCAGCTTGATGCTGACCCAGCAACAGCATCTCTGGAACCTCAG 1026
 Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
 Db 1027 TTGAGCAGCAGCAGCTGGTGTACTTGGCATCCAGGAGCTGTGCCGCGCTGAGCCAG 1086
 Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
 Db 1087 CCGGGCACCACACTGCGGTGCTCTGCTTGGGAGCTGTGAGGTGACCAACAGCGGCTGC 1146
 Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
 Db 1147 AGCAGCTCTGCTGCTCTGCTGCGCAACCCGAGCTCTGCGAGAGCTGAGACCTGAGCAAC 1206
 Qy 1040 AsnAsnPheSerProLysGlyMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
 Db 1207 AACTGTGGGCGACCGCGGCTCTGCGAGCTGTGGGAGCTGGAGCGCGGCTGC 1266
 Qy 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
 Db 1267 GCCTGAGCAGCTGTGCTGTACGACACTACTTGGCGGAGGAGGTGGAGGAGCGCCTG 1326
 Qy 1080 GluGluValGlnLeuLysPro-----ArgValVal 1090
 Db 1327 CAGGCTCTGGAGGAGCAAGCCCGGCTGAGGGTCAATC 1365

RESULT 6

; US-08-910-731-7
 ; Sequence 7, Application US/08910731
 ; Patent No. 5932440
 ; GENERAL INFORMATION:
 ; APPLICANT: SHANTERJEE, DEB K.
 ; APPLICANT: CHANDILYA, HARINI
 ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 ; CITY: WASHINGTON
 ; STATE: D.C.

COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA

US-08-910-731-7

Alignment Scores:

Pred. No.:	2,19e-54	Length:	1371
Score:	612.00	Matches:	152
Percent Similarity:	49.26%	Conservative:	81
Best Local Similarity:	32.14%	Mismatches:	214
Query Match:	8.65%	Indels:	26
DB:	2	Gaps:	5

US-10-066-521-6 (1-1344) x US-08-910-731-7 (1-1371)

Qy	625	AsnLeuAspValLeuAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysValle	644
Db	4	AACCTGGACATT-----CATTCGAGCAGCTGAGCGAGCC	39
Qy	645	ArgValAspValLysGlyPheProArgAspGluSerAlaGluAlaCysProValVal	664
Db	40	CGG-----TGACAGAGCTCCTCCGCTGCTC	66
Qy	665	ProLeuTrp-----MetArgAspLysThrLeuLeuGluGlnTrpGluAsp	680
Db	67	CAGCAGTATGAGTGGTTCAGGCTCGACCTCGCGCTTCACGGAGGAGCACTGCAAGGAC	126
Qy	681	PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerile	700
Db	127	ATCGGTTCCTCCCGGGCAACCCCTCCCTGACCGAGCTCTGCCTCCGACCAACGAG	186
Qy	701	LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle	720
Db	187	CTGGCGGATGCGGGTGCACCTGGTGTGCGAGGCTGCGAGCCCACTGCAAGATC	246
Qy	721	GlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGlyValGlnHisLeuTrpArg	739
Db	247	CAGAAGCTCAGCTGCAGAACTGCTCCCTGACCGAGGGGGCTGCGGGGTCTGCCCCAGC	306
Qy	740	IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGlu	759
Db	307	ACGTGCGCTCCCTGCCCAACGCTGCGGAGGTGCTATCTCAGCGACAACCCACTGGGGGAC	366

Qy	760	GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu	779
Db	367	CGCGGCTGGGCTGCTCTGTGAGGGGCTCTGGACCCCGCAGTCCACCTCGAGAGCTG	426
Qy	780	ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr	799
Db	427	CAGTTGAGTACTGCGGCTGAGCGGCGCAGCTGCGAGCCCTGGGCTCGGTCTCAGG	486
Qy	800	ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal	819
Db	487	GCACGCGGCGCTTGAAGGAGCTCACGTTGAGCAACACACATCGCGGAGCCGCGGCC	546
Qy	820	MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu	839
Db	547	CGGCTGCTGGCCAGGGTCTGGCCGACTCTGCCCTGCCAGCTGGAGAGCTCAGGCTG	606
Qy	840	AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg	859
Db	607	AACCTGCGCTCAGCCAGCCAACTGCAAGACCTGTGCGGAATTGTGGCTCCAGGCC	666
Qy	860	SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu	879
Db	667	TCGCTGCGGAGCTGGCCCTGGGCAGCAACAAGCTGGGTGATGTGGCATGGCGGAGCTG	726
Qy	880	CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis	899
Db	727	TGCCAGGGCTGTCCACCCAGCTCCAGGCTCAGGACCTGTGGATCTGGAGTGTGGC	786
Qy	900	LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr	919
Db	787	ATCACTGCCAAGGGCTGCGGGGATCTGTGCGGTCTCTCAGGGCCAAAGGAGAGCTCA	846
Qy	920	HisLeuSerLeuSerMetAsnProValGluAsnGlyValLysLeuLeuCysGluVal	939
Db	847	GAGCTCAGCTGCGCGCAACGAGCTGGGGATGAGGGTGCCTGCTGTGTGAGACC	906
Qy	940	MetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeuThrAla	959
Db	907	CTGCTGGAACCTGGCTGCCAGCTGGAGTCTGTGGGTGAAGTCTTCAGCTTCACAGCC	966
Qy	960	AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp	979
Db	967	GCCTGTGCCCCCACTTCAGCTCAGTCTGCTGGCCCAAGAGGTTCTCTCTGGAGTAC	1026
Qy	980	LeuThrAspAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln	999
Db	1027	ATAAGCAACACAGCTGGAGGATGCGGGCTGCGGAGCTGTGCCAGGGCTTGGGCCAG	1086
Qy	1000	LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys	1019
Db	1087	CCTGGCTCTGTGCTGGGGTCTCTGGTGGCCGCTGCGATGTGATGTGATGACGAGCTGC	1146
Qy	1020	GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln	1039
Db	1147	AGCAGCTCCCGCAACCTGTGGCCAAACACAGCCCTGCTGAGCTGAGCTGAGCTCAG	1206
Qy	1040	AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer	1059
Db	1207	AACCTGCTGGGAGCGCGGCATCTCTGCACTGTGGAGAGCGTCCGCGAGCCCGGGTGC	1266
Qy	1060	AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu	1079
Db	1267	CTCTGGAGAGCTGGTCTGTACGACATTTACTGGTCTGAGGAGATGGAGGACCGGCTG	1326
Qy	1080	GluGluValGlnLeuLysPro-----ArgValVal	1090
Db	1327	CAGGCCCTGGAGAGGACAAGCCATCCTCCTGAGGCTATC	1365

RESULT 7

US-08-910-731-5
; Sequence 5, Application US/08910731
; Patent No. 5932440

; Sequence 1904, Application US/09833381
; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1904

; LENGTH: 1515

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(1515)

; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1904

Alignment Scores:

Pred. No.: 5,448-38 Length: 1515
Score: 458.00 Matches: 149
Percent Similarity: 45.83% Conservative: 87
Best Local Similarity: 28.93% Mismatches: 221
Query Match: 6.47% Indels: 60
DB: 4 Gaps: 13

US-10-066-521-6 (1-1344) x US-09-833-381-1904 (1-1515)

Qy 4 AspLysSerLeuThrPheSerSerTyrGly-----LeuGlnTrpCysLeuTyrGluLeu 21
Db 60 GACAAGATGGTGCTTCGCGCAGATGGGCTTCAACCTGCGAGGCTCTCTCGAGCAGCTC 119
Qy 22 AspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSerThr 41
Db 120 AGCCAGGATGAGTTCAGCAAGTTCAGATATCTGATCAGCAGCTTCTCTCCCTGGCACACGAG 179
Qy 42 ThrCysSerIleProGlnPheGluIleGluAlaSerValGluCysLeuAlaLeuLeu 61
Db 180 CTCGAGAGATGCTCCCAACAGAGGATAGCAAGGCTGATGGAGCACTGGTAGAATC 239
Qy 62 LeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsn 81
Db 240 CTCACCACTTGTGACAGCTACTGGGTGAGATGGCGAGCTCCAGGCTTTTGAAAG 299
Qy 82 MetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysIleSerGlnAla 101
Db 300 ATGCACCAATGGATCTCTCTGAGAGAGCAAGGATGAAGTCAGAGAGCAGCTTTGAAA 359
Qy 102 MetGluGlnGlu-----GlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 TCCTTTTAATAAAGGAAGCCTCTATCATTAGGATTAACACGGAAGAACGACCACTCTA 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GACATCGACCAATTAAGAA-----CTCTCAAAACAAACAAACAA----- 461
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMet---ThrLysPheAlaGluGlu 154
Db 462 -----GACAAAGACAATAGGTGACAGGTATATATTGAAGACGAGTTTC----- 503
Qy 155 GluAspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 504 -----CGGAGATCTGAAGAGCTGGCTGAGATAGCAAGAGGCTCCAGTTATG 554
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArg----- 188
Db 555 GCTGAGATACAGATCTGATCCATTTAGCAACCCAGGGGTCTTCGCGGCGCCCTTC 614
Qy 189 -----ThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 206

Db 615 TCATACACGGTGGTGTGTATGTCTCTGAGGCGCTTGGGAAACACCGCTGGCCACAGAA 674
Qy 207 IleValLeuCysTrpAlaGlnGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 675 CTAATGTAGACTGGGACAGAGCAACCTCATCCAC---AAATTCAAATATGCGTTCTAC 731
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 732 CTCAGCTGCAGGAGCTCAGCCGCTGGGCGCGTGCAGTTTTCAGAGCTGGTCTTCAGG 791
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 792 GACTGGCTGAATTCAGAGTATGATTCACACATCTTAGCCCAAGCACGGAATCTTG 851
Qy 267 PheIleIleAspGlyPheAspAspLeuGlySer-----ValLeuAsnAsnAspThr 283
Db 852 TTGCTGATTACGGCTTTGATGAGCTGGGAGCGGACCTGGGGCGCTGATCGAGGAC--- 908
Qy 284 LysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeu 303
Db 909 ---ATCTCGGGGACTGGGAGAGAGAGCGCGTGCCTCTCTCTGGGGAGTTTCTG 965
Qy 304 ArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGlu 323
Db 966 AACAGGCTGATGTATCCCAAGGCGCGCTCTGTGTACACACGCGCCAGGCGCCCTGAGG 1025
Qy 324 LysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGlu 343
Db 1026 GACCTCCGATCTCTGGCGAGGAGCGGATCTATACATAGGCTGGAGGCTTCTCTGGAGAG 1085
Qy 344 GlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeu 363
Db 1086 GACAAGAGGCGCTATTCTCTGAGACACTTTGGAGACGAGGACCAAGCCATCGCGCTTT 1145
Qy 364 ArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySer 383
Db 1146 GAGCTAATAGGAGCAACGCGGCCCTGCTCCAGCTGGGCTCGGCCCGCGGCTGTGCTGG 1205
Qy 384 LeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsn 403
Db 1206 ATCGTGTGACAGCTCTGAAGCTGACATGAGAGGAGGAGGAGGACCGGCTCCCACTGC 1265
Qy 404 GlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro----- 420
Db 1266 CTCACCGCAGCGGGCTGTT-CTTGGCTTCTCTGAGCGCGGTTCGCGAGGCGCACAG 1324
Qy 421 ---ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1325 NTGCGGGCGCGCTCGACGCTGAGCCTCTGCGCGCGGCGGCTGTG----- 1372
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
Db 1373 -----GGGCAATGTCTCGTNCACCGAGAGGAGCTG 1405
Qy 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1406 GNAAGGCTCGGGTGCAGGAGTCCGACNTCCGNTGTTTCTTGTGACGAGAGANATCTCG- 1464
Qy 480 ProAsp-----SerHisCysGluGluTyrTyrThrPhePhe 491
Db 1465 CCAGACAGAGCTCCAAAGGCTGTCGCTTATCCANTCAGNTTC 1509

RESULT 9

US-09-833-381-1905

; Sequence 1905, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

;; PRIOR APPLICATION NUMBER: 09/516,448
;; PRIOR FILING DATE: 2000-02-29
;; NUMBER OF SEQ ID NOS: 2050
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1905
;; LENGTH: 1515
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(1515)
;; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1905

Alignment Scores:
Pred. No.: 5,44e-38 Length: 1515
Score: 458.00 Matches: 149
Percent Similarity: 45.83% Conservative: 87
Best Local Similarity: 28.93% Mismatches: 221
Query Match: 6.47% Indels: 60
DB: 4 Gaps: 13

US-10-066-521-6 (1-1344) x US-09-833-381-1905 (1-1515)

Qy 4 AsplysSerLeuThrPheSerSerTyrGly-----LeuGlnTrpCysLeuTyrGluLeu 21
Db 60 GACAAAGATGGTCTTCGGCGCAGATGGCTTCAACCTGCAGGCTCTCTGGACGACCTC 119
Qy 22 AsplysGluGluPheGlnThrPheLysGluLeuLeuTyrLysSerSerGluSerThr 41
Db 120 AGCCAGGATGATGTCAGCAAGTTCAGATGATCTGATCAGCAGCTTCTCCCTGGCACAG 179
Qy 42 ThrCysSerIleProGlnPheGluLeuGluAsnAlaValGluCysLeuAlaLeuLeu 61
Db 180 CTCAGAGATGCCCCCAAGAGAGGTAGCAAGGCTGATGGGAAGCAACTGGTAGAAATC 239
Qy 62 LeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsn 81
Db 240 CTCACCCCAATGTGACAGTACTGGTGAGATGGCGAGCTTCCAGGCTCTTGAAGAAG 299
Qy 82 MetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysLysIleSerGlnAla 101
Db 300 ATGCACCAAGATGGATCTGTCTGAGAGAGCAAGGATGAAGTCAGAGAAGCAGCTTTGAAA 359
Qy 102 MetGluGlnGlu-----GlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 TCCTTTAATAAAGGAAGCCCTCTATCATATTAGGGATAACACGGAAGAACGACCACTCTA 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 420 GACATCGACCAATTAAGAA-----CTCTTCAAAACAAAAACAA----- 461
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMet---ThrLysPheAlaGluGlu 154
Db 462 -----GACAAAGCAATAGGTGCGAGGTATATATTGAAGCAAGTTC----- 503
Qy 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 504 -----CGGAGATGTGTGAAGAGTGGCGCTGGAGATAGCAAGAGGTCACAGTTATG 554
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArg----- 188
Db 555 GCTCAGAGATACAGATGCTGATCCCATTTAGCAACCCAGGGTGTTCGGGGCCCTTC 614
Qy 189 -----ThrValValLeuHisGlyLysSerGlyLeuGlyLysSerAlaLeuAlaArgArg 206
Db 615 TCATACACGGTGGTGTATGTCCTGAGGCTTGGGAAACCAACGCTGGGCCAGAAA 674
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 675 CTAATGTAGATGGGAGAGCAACCTCATCCAC---AAATTCAATATGCGTTCTAC 731
Qy 227 LeuProValArgGluMetGlnArgLysGluSerSerValThrGluPheIleSerArg 246

Db 732 CTCAGTGCAGGAGCTCAGCCCTGGGCCCTGGAGTTTTCAGAGCTGGTCTTCAGG 791
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 792 GACTGGCTGAATTGCAGGATGACATTCACACATCTTAGCCCCAAGCAGGAAATCTTG 851
Qy 267 PheIleIleAspGlyPheAspLeuGlySer-----ValLeuAsnAsnAspThr 283
Db 852 TTCGTGATTGACGGCTTTGATGAGCTGGAGCGCCACCTGGGGCGGTGATCGAGGAC-- 908
Qy 284 LysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArgSerLeuLeu 303
Db 909 ---ATCTCGGGGACTGGGAGAAAGAACCCCGTCCCTCTCTGGGGATTTGCTG 965
Qy 304 ArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGlu 323
Db 966 AACAGGTTGATGTTACCAAGCGCCCTGCTGTGTACACAGCGGCCAGGGCCCTGAGG 1025
Qy 324 LysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyLysSerGlyGlu 343
Db 1026 GACCTCCGATCTCGCGGAGGAGCGCATCATAAAGGTGGAGGGCTTCTCTGGAGGAG 1085
Qy 344 GlnArgIleHisLeuLeuGluArgGlyLleGlyGluHisGlnLysThrGlnGlyLeu 363
Db 1086 GACAAGAGGCTTATTTCTGAGACACTTTGGAGAGCAGGACCAAGCCATGCGTGCCTT 1145
Qy 364 ArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySer 383
Db 1146 GAGCTAATGAGGAGCAACGCGNCCCTGTTCCAGCTGGGCTCGGCCCGCGGTGTCTG 1205
Qy 384 LeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsn 403
Db 1206 ATCGTGTGACGACTCTGAAGCTGCAGATGNAGAAAGGGAGAGCCCGTCCCGACCTGC 1265
Qy 404 GlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro----- 420
Db 1266 CTCACCCGACGGGGCTGTT-CTTGGCTTCTCTGACGCCGCTTCCCGAGCGGCACAG 1324
Qy 421 ---ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1325 NTGCGGGCGCGCTGCGAGCTGAGGCTCTTGGCGCGCAGGGCTGTG----- 1372
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
Db 1373 -----GGGCAGATGTCGCTGTNCCACCGAGAGGACCTG 1405
Qy 460 MetValGlnGlyLeuGlySerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1406 GNAAGCTCGGGTGCAGGAGTCCGACNTCCGNTGTTCTCTGACGAGGAGATCTCCG- 1464
Qy 480 ProAsp-----SerHisCysGluGluTyrTyrThrPhePhe 491
Db 1465 CCAGACAGAGTCTCCAAAGGCTGTNGCCTTATCCANTCAGNTTC 1509

RESULT 10

US-09-833-381-1976
; Sequence 1976, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1976
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens

Db 2388 CTGGCAACAACAACTCAATGACTATGCGGTGAC-----GAGCTGACCGCTTGTCTT 2441
Qy 668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db 2442 AGCGT----- 2447
Qy 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLys 707
Db 2448 -----CTCAGCGTTATCATGACTCATCGCTCAACACAGATCATCCGACACCGGGGTGAAG 2498
Qy 708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet---PheArg 726
Db 2499 GTGCTATGTAGGAAGT-----ACCAAGTATAGATGTGAGCTTCTCGGGTTTATAC 2552
Qy 727 AsnAlaGlnIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Db 2553 AACCAACAGATACTGATATCGGACCCAGGTATGTGGCCCAATCTCTGGATGATGACGA 2612
Qy 746 AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAla 765
Db 2613 GGCCTCAAGCACCTTAAACTAGGGAACCAAGATAACAAGTGAAGCGGGAAGTGTGTG 2672
Qy 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGly 785
Db 2673 GCTTGGCTGTGAGAAC----- 2690
Qy 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
Db 2691 -----AGCACCTCCATCGTT 2705
Qy 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825
Db 2706 GATGTTGGGATGTGGGTAAATCAGATTGGACGACGAAGGGCAAGCCCTTCGCAGAGCA 2765
Qy 826 LeuArgValSerGlnCysAlaLeuGlnLysIleLeuLeuGluAspCysGlyIleThrAla 845
Db 2766 TTGAAG---GACCACCCCGCTGACCACTCTCAGTCTTGATCTCAATGGCATCTCCG 2822
Qy 846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db 2823 GAGGAGGAGAGAGCTTCCAGCGCTTGAAGCAGACACACACACATGACATGATCTGG 2882
Qy 866 LeuSerAsnSerLeuGlyAsnGlyValAsnLeuLeuCysArgSerMetArgLeu 885
Db 2883 CTGACCAAAATGAATTAATGATGAGTCTGACAGTCTTGGCTGAGATGCTCAGAGTG 2942
Qy 886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
Db 2943 ---AACACAGCGCTACGGCATTTATGGCTGATCCAGAAATCCATCAGACCAAGGGACA 2999
Qy 906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
Db 3000 CGCAGCTGGCGAGGCACTGCAGNAGAACACACCCATACAGAGATTTGTCTCAATGGA 3059
Qy 926 AsnProValGlu-----AspAsnGlyValLysLeuLeuCysGlu 938
Db 3060 AACTTGATTAAAGCCGAGGCGCAAGTCTTCGAGAAATGAGAGAAATCATCTGCTT- 3118
Qy 939 ValMetArgGluProSerCysHisLeuGlnAspLeu-----GluLeuValLysCysHis 956
Db 3119 CTGACGGACGCTCTGGG-----CAGGATCTTTGTCTAGGTTGCTCTCAGTCAC 3169
Qy 957 LeuThrAlaAlaCys-----CysGluSerLeuSerCysVal 968
Db 3170 AG-ACAGACCTGTGAGTCAGCAGGAGTAGCAGGATGTGTGACGCGCTGCAGCAAGGTG 3228
Qy 969 IleSerArgSerArgHisLeuLysSer 977
Db 3229 CCTGTGAGGAGCCACACCTCCACAGT 3255

RESULT 12

US-09-207-359B-42

; Sequence 42, Application US/09207359B

; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; CURRENT FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 4141
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (261)...(3119)
US-09-207-359B-42

Alignment Scores:
Pred. No.: 7,36e-29 Length: 4141
Score: 379.50 Matches: 253
Percent Similarity: 37.01% Conservative: 150
Best Local Similarity: 23.23% Mismatches: 416
Query Match: 5.36% Indels: 273
DB: 4 Gaps: 46

US-10-066-521-6 (1-1344) x US-09-207-359B-42 (1-4141)

Qy 6 SerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGlu 25
Db 451 AGCTGACAAAGGTCGGAAGATCCTTGACCTGTGC----- 486
Qy 26 PheGlnThrPheLysGluLeuLeuLysSerSerGluSerThrThrCysSerIle 45
Db 487 -----AGACCAAGGGCAGGAGGTCTGAGTCTTCTCTAGTGTGAG- 533
Qy 46 ProGlnPheGluLeuGluAsnAlaAsnValGluCysLeuAlaLeuLeuHisGluTyr 65
Db 534 -----CAGCTGGAGGATGCTTACGTGAC-----CTCAGGCTG- 566
Qy 66 TyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsnMetAsnLeuArg 85
Db 567 -----TGGCTCTCAGAAATTTGGCTTCTCCCTTCCAGCTCATTTCCG 608
Qy 86 ThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGlnAlaMetGluGlnGlu 105
Db 609 ACCAAACTATCTCAATACTGACCCAGTAAGCAGGTATACCCACAGCTGCGACACCAA 668
Qy 106 GlyAlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGly 125
Db 668 ----- 668
Qy 126 AlaThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
Db 669 -----CTGGCGCGAC-----TCC 683
Qy 146 HisValMetThrLysPheAlaGluGluAspValArgSerPheGluAsnThrAla 165
Db 684 AAGTTCACTGTCTGTACGCCAGAGAGGAGACTG-----CTGCTGGAGGAGACC- 734
Qy 166 AlaAspTrpProGluMetGlnThrLeuAlaGly-----AlaPheAspSerAspArgTrp 183
Db 735 -----TATATGGACACACTCATGGGCTGTGTAGGCTTCAACATGAACACCTG 782
Qy 184 GlyPhe-----Arg 186
Db 783 GGCAGCGCTAGGAGCGCTGGATTGCTCTGACACACAGTACGGGCGTCTCAACAGAGCAT 842

1302 ---CTGCTCGGGCTTCTCCCAAGT-----CACCTGGCGGCTATGCCGCGCG 1349
Qy LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet 367
Db ATGTTCCCGAGCGCACAGCGCAGGAGCATCTG-----CTGCAGCAGCTGGAT 1397
Qy AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal 387
Db GCCAACCCCAACCTCTGCAGCTGTGCGGGTGGCGTCTTCTGTGTGATCATCTCCGT 1457
Qy AlaLeuGln---LeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db TGTTCACAGCATCTCCAGACGGTCTTCGAGGCTCTCTTCACAGTTGCCGAGCTGTGCT 1517
Qy ThrGlyLeuHisAlaAlaPheVal-----PheHisGlnLeuThrPro 420
Db GTGACCTTGACCGATGTCTTCTGTGTCACAGTGCATCTGAACAGGCGCGAGGCC 1577
Qy ArgGlyValValArgAtqCysLeuAsnLeuGluGluArgValVal----- 435
Db AGCAGCTGGTGCAGCGCAACACGGCGAGCCCGCGGAAACCTACGTGCAGGTGGCGC 1637
Qy ---LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
Db ACGCTGCATCGCTGGAGAGGTGGCTCACCGAGGCGCCGACAGAGCCCTTTTGTGTT 1697
Qy AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
Db GCGCAGGAGGAGGTGTCAGGCTCGAAAGCTGCAGGAAGGAGATCTGCAGCTGGGCTCTCGT 1757
Qy MetAsnIleLeuLeuProAspSerHisCysGlu-----GluTyrTyrThrPhePheHis 492
Db GCTTTGCCCGATGTGGGCCCTGAGCAGGCGCAGTCTTACGAATTTTCCAC 1811
Qy LeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGlu 512
Db CTACGCTCCAGGCTTCTTCCCGCTTCTTCTGTGTAGCAGATGCACAAAGTGCAGCAC 1871
Qy ProAlaLeuCysProLeuTyrValGlnLysThrLysArgSerMetGluLeuLysGlnAla 532
Db CGGAGTTGTGTGAGTCTTTTCGAGAATGGACGTCTCTCTGGA---GAGGCAACAAGCTCG 1928
Qy GlyPheHisIle----- 536
Db TCCTGCCATTTCTCTTCTCTTCCATGTCCTGGGCGCAGAGCGGTTGGGCGCT 1988
Qy ---HisSerLeuTrpMetLysArgPheLeuPheGlyLeuVal 549
Db GATCCTTTTCAGGAACAAGATCACTTCCAGTTTCCACCACTCTTCTGTGTGGGCTACTG 2048
Qy SerGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
Db GCCAAAGCCGACAGAACTCTTTCGGCAGCTG-----GTGCCCAAGGCTATC 2096
Qy ---LysGlnLysLeuLeuHisTrpVal-----SerLeu 579
Db CTGAGGAGGAAGCGAAGGCCCTG---TGGCTCACCTGTTTGTAGCTGCGCTCTCTAC 2153
Qy LeuGlyGlnGlnProAsnAlaThrThrProGly-----Asp 591
Db TTGAAGAGCCTACCTCGGCTCCAGTCTGAGGCTTTAACCAAGGTGCATGCCATGCCACA 2213
Qy ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db TTTCTGTGTGATGCTCGCGTGCATCTATGACGCGCAGCCAGAGGTGGGCGCGCTCGCC 2273
Qy LeuAsnSerPheGlnGluValTrpLeuProIle-----AsnGlnAsnLeuAsp 627
Db GCCAGGGGCATCAGTGGGAGTACTTCAAGTGGCTTTTGGCAACGCTTGTCTTGGCGGAC 2333
Qy LeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAsp 647

Db 2334 TGCAGCGCCCTGCTTCTGTCCTGCATCAC-----TTCCACAGGCGAGCTGGCCCTAGAC 2387
Qy ValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrp 667
Db CTGCACAACAACAACCTCAATGACTATATGGCGTGCAG-----GAGTGCAGCCTTGCTTT 2441
Qy MetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db AGCCGT----- 2447
Qy HisProHisLeuArgGlnLeuAspLeuLysSerSerIleLeuThrGluArgAlaMetLys 707
Db ---CTACGGTTATCAGACTCAGCTCAGCATCAACAGATCACCCACAGCGGGTGAAG 2498
Qy ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet---PheArg 726
Db GTGCTATGTGAGAACTG-----ACCAAGTATAAGATCGTGACGTTCTCTGGGTTTATAC 2552
Qy AsnAlaGlnIleThrPro---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Db ---GlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Qy AsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAla 765
Db GGCCTCAAGCACCTTAACTAGGGAATAACAGAAATCAAGTGCAGGCGCGGAAGTGTGTG 2672
Qy CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGly 785
Db GCTTTGGCTGTGAAGAAC----- 2690
Qy LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
Db -----AGCCTCCATCGTT 2705
Qy SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825
Db GATGTTGGGATGTGGGTAAATCAGATTGGAGAGGAAGGCAAGGCTTCGCAGAGGCA 2765
Qy LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAla 845
Db TTGAAG---GACCACCCCGACCTCTCAGTCTTTCATTCATTCATTCATCTCTCCG 2822
Qy ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db GAGGAGGAGGAGGCTTGGCGAGGCCCTGAGCAGCAACACCACTGCAGTAATCTGG 2882
Qy LeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeu 885
Db CTGACCAAAATGAACCTTAATGATGAGTCTGCAGAGTGTCTGCTGAGATGCTGAGAGTG 2942
Qy ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
Db ---AACACAGCGTACCGCATTTTATGGCTGATCCAGAATCGCATCACAGCAAGGGGACA 2999
Qy GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
Db GCCCAGCTGGCGAGGCACTGCAGAGAACAACACAGCCATACAGAGATTTGTCTCAATGGA 3059
Qy AsnProValGlu-----AspAsnGlyValLysLeuLeuCysGlu 938
Db AACTTGATTAAAGCCCGAGGAGGCAAAAGTCTTCAGAAATGAGAAGAGAAATCATCTGCTT 3118
Qy ValMetArgGluProSerCysHisLeuGlnAspLeu-----GluLeuValLysCysHis 956
Db CTGACGAGCGCTCTCGG-----CAGGATCTTTGTCTTAGGTTGTCTCTCACTGCTCAC 3169
Qy LeuThrAlaAlaCys-----CysGluSerLeuSerCysVal 968
Db AG-ACAGACTGTGCAGTGCAGCGGTAGCAGGATGCTGTGCAGCGCCTGCAGCAAGGTG 3228
Qy IleSerArgSerArgHisLeuLysSer 977
Db CCTGTGAGGAGGCCACACCTCCACACT 3255

```

RESULT 14
US-09-865-364-42
; Sequence 42, Application US/09865364
; Patent No. 6613521
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; ; SEQUENCE 42, PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/865,364
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 4141
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (261) ... (3119)
US-09-865-364-42

```

Alignment Scores:	
Pred. No.:	7.36e-29
Score:	379.50
Percent Similarity:	37.01%
Best Local Similarity:	23.23%
Query Match:	5.36%
DB:	4
	Gaps: 46
	Indels: 273
	Mismatches: 416
	Conservative: 150
	Matches: 253
	Length: 4141

US-10-066-521-6 (1-1344) x US-09-865-364-42 (1-4141)

Qy	6	SerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGlu	25
Db	451	ACGCTGCAAGGTCGGAAGAATCCTTGACCTGGTGTC	486
Qy	26	PheGlnThrPheLysGluLeuLeuLysLysLysSerSerGluSerThrThrCysSerIle	45
Db	487	-----AGACGAAAGCGAGGAGTGCTGAGTTCTTCCTCTACGTGCTCCAG- 533	
Qy	46	ProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeuLeuLeuHisGluTyr	65
Db	534	-----CAGCTGGAGGATGCTTACGTGGAC--CTCAGGCTG----- 566	
Qy	66	TyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsnMetAsnLeuArg	85
Db	567	-----TGCTCTCTCAAAATTTGGCTTCTCCCTTCCCGAGCTCATTCGG 608	
Qy	86	ThrLeuSerGluLysAlaArgAspMetLysLysIleSerGlnAlaMetGluGlnGlu	105
Db	609	ACCAAACTATCGTCAATCTGACCCAGTAGCAGGTATACCCAAACAGCTGGCACCAAA 668	
Qy	106	GlyAlaThrAlaAlaGluThrGluGluGlnGluLysSerGlnAlaMetGluGlnGluGly	125
Db	668	----- 668	
Qy	126	AlaThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer	145
Db	669	-----CTGGGCGCGCAC-----TCG 683	
Qy	146	HisValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAsnThrAla	165
Db	684	AGTTCATGCTGTGTAGCCCCAGAGGAGGACCTG-----CTGCTGGAGGAGACC--- 734	
Qy	166	AlaAspTrpProGluMetGlnThrLeuAlaGly-----AlaPheAspSerAspArgTrp	183

Db	735	-----TATATGGACACACTCATCGGGCTGGTAGCTTCAACATGAAACCTG	782
Qy	184	Glype	186
Db	783	GGCAGCCCTAGAGGCGCTGGATTGCCTGCGACACACAGTAGTCGGCGCTCCTCAACAGCAT	842
Qy	187	ProArgThrValValLeuHisGlyIysSerGlyIleGlyIysSerAlaLeuAlaArgArg	206
Db	843	GGCGAGACTCTCTTCGTGTTCGGGAGCGCGGAGTGGCAAGTCCATGCTGCTGCACAGG	902
Qy	207	IleValLeuCysTirpAlaGInGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe	226
Db	903	TTGCAGAGCCCTCTGGGGCTCAGGAGGTG---ACCTCCACAGCCAAATTCCTCTTCCAC	959
Qy	227	LeuProValArgGluMetGlnArgLysLysGluSer-----SerValThrGluPhe	243
Db	960	TTCCGCTGCCGATGTTTCAGCTGCTTCAAGGAGAGGACATGCTGAGTCTGCAGGACCTG	1019
Qy	244	IleSerArgGluTrp-----ProAspSerGlnAlaProValThrGluIleMet	258
Db	1020	CTCTTCAAGCATTTCTGCTACCCGGAGCAGACCCCGAGAGGTGTTCTCTCTTCTGCTG	1079
Qy	260	SerArgProGluArgLeuLeuPheIleAspGlyPheAspPheLeuGlySer-----	277
Db	1080	CGCTTTCCACACAGCGCTCTTCACTTTTGACGGCTGTGATGAGTGCACATCAGATTC	1139
Qy	278	ValLeuAsnAsnAspThrLysLeuCysLysAspTirpAlaGluLysGlnProProPheThr	297
Db	1140	GACCTGAGCGCGTCCGGATAGCTGTCGCCCTGG---GAGCGGCTCACCTCTGGTC	1196
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1197	CTGCTGCTAACCTCTCAAGTGGGAGCTGCTCAAGGGTCCCGCAAAATGCTCACTGCT	1256
Qy	318	Arg-----AspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr	334
Db	1257	CGCACAGCGCTGAGGTCCCCCGCAGCTCTCGCAAAAGGTG-----	1301
Qy	335	LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeu-----	348
Db	1302	---CTGCTCGGGCTCTCCCCAAGT-----CACCTGGCGGCTATGCCGCGG	1349
Qy	349	---LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet	367
Db	1350	ATGTTCCCCGAGGCGACAGCGCAGGACATCTG-----CTGCAGCAGCTGGAT	1397
Qy	368	AsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal	387
Db	1398	GCCAAACCCCACTCTGCAGCCTGTGGGGTGCCGCTCTCTGTGTGATCATCTCCGT	1457
Qy	388	AlaLeuGln--LeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu	406
Db	1458	TGTTTCCAGCATTCACAGCGTCTTCGAGGGCTCTCTTCACAGTTGCCGGACTGTGCT	1517
Qy	407	ThrGlyLeuHisAlaAlaPheVal-----PheHisGlnLeuThrPro	420
Db	1518	GTGACCTGACCGATGCTTTTCTGCTGGTCACTGAGTGCATCTGAACAGCGCCGACGCC	1577
Qy	421	ArgGlyValValArgCysLeuAsnLeuGluGluArgVal-----	435
Db	1578	AGCAGCCTGTGTGGGCGCAACACGCGCAGCCCGCGGAACCCCTCATGTCAGCGCTGC	1637
Qy	436	---LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe	454
Db	1638	ACGTGTCATCGCTGGGAGGTGGCTTCACCGAGGACCGACGAAGACCTCTTTGTGTTT	1697
Qy	455	AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis	474
Db	1698	GGCCAGGAGGATGTCAGGGCTCGAAGCTGCAGAGGAGATCTGCAGCTGGGCTTCCGT	1757
Qy	475	MetAsnIleLeuLeuProAspSerHisCysGlu-----GluTyrTyrThrPhePheHis	492
Db	1758	CGG-----CCTTTGGCCGATGTGGGCGCTTCAGCAGGGCGCAGTCTTACGAATTTTCCAC	1811

189	ThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgAlaVal	208
833	ACCATCTTCATCTCGGTGATGTGGGGTGGCAAGTCCATGTCTGTACAGCGCGTGCAG	892
209	LeuCysTTPAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPheLeuPro	228
893	AGCCTCTGGGCCACGGCGCGGTAGACGACAGGGGTC--AAATTCCTTCTTCCACTTTCGC	949
229	ValArgGluMetGlnArgLysLysGluSerSer-----ValThrGluPheIleSer	245
950	TGCGCGATGTTACAGCTGCTTCAGAGAAAGTGACAGGCTGTGTGACGAGACCTGCTCTTC	1009
246	ArgGluTTP-----ProAspSerGlnAlaProValThrGluIleMetSerArg	261
1010	AAGCACTACTGCTTACCACAGACGGGACCCCGAGGAGGTGTTTGCTTCTGCTCGCGCTTC	1069
262	ProGluArgLeuLeuPheIleIleAspGlyPheAspPheGlySerValLeuAsnAsn	281
1070	CCCCAGTGGCCCTCTTCACCTTCGATGGCTCGACAGCTGCACCTCGGACTTG-----	1123
282	AspThrLysLeuCysLysAsp-----TTPAlaGluLysGlnProProPheThr	297
1124	GACCTGAGCGCGTGTGACAGCTCTGCGCCCTGG--GAGCCTGCCACCCCTGGTC	1180
298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
1181	TTGCTGGCCAACTGCTCAGTGGGAAGCTGCTCAAGGGGGCTAGCAAGCTGCTCACAGCC	1240
318	Arg-----AspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr	334
1241	CGCACAGCATCAGAGTCCGCGCCAGTTCTCTCGGAAGAAGGTG-----	1285
335	LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeu-----	348
1286	---CTTCTCGGGGCTTCTCCCCAGC-----CACCTGGCGGCTATGCCAGGAGG	1333
349	---LeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMet	367
1334	ATGTTCCCCGAGCGGCGCTGCAGGACCGCTGCTGAGCCAG-----CTGGAG	1381
368	AsnAsnArgGlnLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVal	387
1382	GCCAAACCCAACTCTGCAGCCTGTCTGTGTGCC-----CTCTTCTGTCTGG	1429
388	AlaLeu-----GlnLeuGln	392
1430	ATCATCTTCGGTGTCTTCAGCAGCTTCCGTGTCCTTTGAAGGCTCACACAGCTGCC	1489
393	AspValValGlyGlyLysValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAla	412
1490	GACTGC-----ACGATGACCTGCAGATGCTTCTCTCTCTG	1525
413	PheVal-----PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsn	429
1526	GTCACCTGAGGTCCATCTCTGAACAGGATGCAGCGACCTGTGTGACGCG-----AAC	1579
430	LeuGluGlu-----ArgValValLeuLysArgPheCysArg	441
1580	ACAGCAGCCAGTGGAGACCTTCCAGCGCGCGGACACTCTGTGCTCGTGGGGCAG	1639
442	MetAlaValGluGlyValTTPAsnArgLysSerValPheAspGlyAspLeuMetVal	461
1640	GTGCCCCACGGGGCATGGAGAGAGAGCTCTTGTCTTCACCCAGGAGGAGGTGCAGGCC	1699
462	GlnGlyLeuGlyGluSerGlu-----LeuArgAlaLeuPheHisMetAsn	476
1700	TCCGGGCTGCAGGAGAGAGACATGCAGCTGGGCTTCTCGGGGCTTTTCCGAGCTGGC	1759
477	IleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGln	496
1760	-----CCCCGGGGTGAC---CACAGTCTCTATGATGTTTTTTCACCTCACCTCCAG	1807

Qy	497	AspPheCysAlaIalLeuTyrTyrValLeuGlu	507
Db	1808	GCCTTCTTTACAGCCTTCTTCTCGTCTGGACGACACAGGGTGGGCACCTCAGAGCTGCTC	1867
Qy	508	-----GlyLeuGluIle-GluProAlaLeu	515
Db	1868	AGGTTCTTCAGAGTGGATGCCCCCTGCGGGGGCAGCAGCAGCTCTGCTATCTCTCC	1927
Qy	516	-----CysProLeuTyrValGluLeuThrLysArgSerMetGluLe	529
Db	1928	TTCTCCCGTTCCAGTGCCTGCAGGGCAGTGGTCCGGCGCGGAAGACCTCTTCAAGAAC	1987
Qy	529	uLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuVa	549
Db	1988	-AAGGAT-----CACTTCCAGTTTCCACCAACTCTTCTGTCGGGGCTGT	2031
Qy	549	lSerGluAspValArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVa	569
Db	2032	GTCAAAGCCAAACAGAAACTCTGCGGCATCTGGTGCCCGCG---GCAGCCCTTAGGAG	2088
Qy	569	lLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsn-----AlaTh	587
Db	2089	AAAGCGCAAGCGCCTG--TGGGCACACCTGTGTTTCCAGCTCGCGGGCTACCTGAAGAG	2145
Qy	587	rThrProGlyAspThrLeuAspAlaPheHis-----	597
Db	2146	CCTGCCCGGTTTCAGGTGCAAAAGTTTCAACACAGTGCAGGCCCATGCCACGTTTCATCTG	2205
Qy	598	-----CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSe	614
Db	2206	GATGTCGCTGTCATCTACGAGACACAGCCAGAGGTGGGCAGCTGGCGCCAGGGG	2265
Qy	614	rPheGlnGluValTrpLeuProIleAsnGln-----AsnLeuAspLeuIleAl	630
Db	2266	CATCTGGCCCACTACTCTCAAGCTGACTACTGCAAGCCCTGCTCGGCCGACTCAGCGC	2325
Qy	630	aSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGl	650
Db	2326	CCTCTCTCTGCTCTGCATCACTTCCCT-----AAGCGGTGGCCCTAGACCTAGACAA	2379
Qy	650	yllePheProArgAspIleSerAlaGluAlaCysProValValProLeuTrpMetArgAs	670
Db	2380	CAACAATCTCAACGACTACCGCGTGGG-----GAGCTGAGCCCTGCTTTCAGCCCG--	2431
Qy	670	plysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHi	690
Db	2431	-----	2431
Qy	690	sLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCy	710
Db	2432	-CTCAGCTGTTCTCAGACTCAGCGTAAACACAGATCACTGAGCGTGGGGTAAAGTGCTTAAG	2490
Qy	710	sAlaLysLeuArgHisProThrCysLysIleGlnThr--LeuMetPheArgAsnAlaGl	729
Db	2491	CGAAGAGCTG-----ACCAAAATACAAAATGTGACCTATTTTGGGTTTATACAAACACCA	2544
Qy	729	nileThrPro--GlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuAr	748
Db	2545	GATCACCAGATGTCGAGCCAGGTACGTACCCAAAATCTTGATGAATGCAAGAGCCTCAC	2604
Qy	748	gSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAl	768
Db	2605	GCATCTTAAACTGGGAAAAAACAATAACAGTGAAGGAGGAAGTATCTCGCCCTGGC	2664
Qy	768	aLeuLysPheIleProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHi	788
Db	2665	TGTGAAGAACAGCAGCAA-----	2680
Qy	788	sAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSe	808
Db	2681	-----TCAATCTGAGGTGG	2697
Qy	808	rLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVa	828

[illegible]

Search completed: July 16, 2004, 06:26:42
Job time : 280 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 15:25:53 ; Search time 7396 Seconds
(without alignments)
5426.551 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 7074
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10066521/runat_13072004_122212_11091/app_query.fasta_1.1543
-DB=EST -QFMT=tastap -SURFIX=rest -MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -List=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521 @CGN 1.1 5759 @runat_13072004_122212_11091 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fut:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2653	37.5	3475	11	AK087774	AK087774 Mus muscu
2	1422	20.1	3359	11	BC021272	BC021272 Homo sapi
3	1422	20.1	3360	11	BC012789	BC012789 Homo sapi
4	1217	17.2	791	14	CK000513	CK000513 AGENCOURT
5	1196.5	16.9	3577	11	AK087843	AK087843 Mus muscu
6	1182.5	16.7	3218	11	AK054378	AK054378 Mus muscu
7	1043.5	14.8	1913	11	AK016782	AK016782 Mus muscu
8	1008	14.2	3793	11	BC036506	BC036506 Homo sapi
9	962	13.6	3275	11	AK054426	AK054426 Mus muscu
10	940	13.3	3400	11	AK054264	AK054264 Mus muscu
11	687	9.7	1119	12	BM454123	BM454123 AGENCOURT
12	669.5	9.5	814	14	CB228956	CB228956 AGENCOURT
13	669.5	9.5	2730	11	AK039531	AK039531 Mus muscu
14	643.5	9.1	1371	29	AY407369	AY407369 Mus muscu
15	636	9.0	671	10	BB555431	BB555431 BB555431
16	628	8.9	684	13	BU630481	BU630481 UI-H-FL0-
17	626	8.8	693	13	BU634350	BU634350 UI-H-FL1-
18	613.5	8.7	2218	11	AF054176	AF054176 Homo sapi
19	608.5	8.6	1371	29	AY407367	AY407367 Homo sapi
20	586.5	8.3	1922	11	AK049352	AK049352 Mus muscu
21	559	7.9	579	12	BM228619	BM228619 K0264H01-
22	558.5	7.9	1313	14	CF110534	CF110534 Shultromi
23	553	7.8	666	9	AV367637	AV367637 AV367637
24	548	7.7	642	13	BU618831	BU618831 UI-H-PH1-
25	543	7.7	565	12	BM226108	BM226108 K0222C06-
26	534	7.5	706	10	BB624558	BB624558 BB624558
27	517.5	7.3	1434	11	AK014932	AK014932 Mus muscu
28	507.5	7.2	1160	9	AF522024	AF522024 AF522024
29	486.5	6.9	589	14	CA559886	CA559886 K0264H01-
30	485	6.9	511	12	BM230035	BM230035 K0288H05-
31	485	6.9	513	12	BM227869	BM227869 K0250F09-
32	483	6.8	519	10	BG071729	BG071729 H3102B07-
33	483	6.8	519	12	BG084574	BG084574 H3102B07-
34	482	6.8	515	12	BM227562	BM227562 K0245F10-
35	481	6.8	512	12	BM226240	BM226240 K0225A02-
36	481	6.8	901	14	CD515309	CD515309 AGENCOURT
37	475.5	6.7	583	14	CA559979	CA559979 K0266B01-
38	473.5	6.7	970	12	BG696072	BG696072 60268033
39	473	6.7	950	13	BQ951319	BQ951319 AGENCOURT
40	470	6.6	898	13	BUI13197	BUI13197 603117379
41	469.5	6.6	593	14	CA558845	CA558845 K0250F09-
42	459	6.5	1246	29	AY407368	AY407368 Pan trogl
43	454.5	6.4	643	14	CF913803	CF913803 B0954D05-
44	449	6.3	1003	12	BM558396	BM558396 AGENCOURT
45	447.5	6.3	636	14	CF914959	CF914959 B0971H02-

ALIGNMENTS

RESULT 1
AK087774
LOCUS
DEFINITION Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330019P16 product:mouse WATER protein (maternal-antigen-that-embryos-require) protein, full insert sequence.
ACCESSION AK087774
VERSION AK087774.1 GI:26104500
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 731 AAAGTGGACAGAGAGAGAGTGGTGGCAGAGCTGATTTGCTAAAGGAGTGTCCAGAGCTCCT 790
Qy 252 InAlaProValThrGluLeuMetSerArgProGluArgLeuLeuPheIleLeuLeuGlyP 272
Db 791 GGGATCTAGTACAAAGATCATGCTCCACACGAAAGACTCTGGTTGTCATAGATGGCT 850
Qy 272 heAspAspLeuGlySerValLeu---AsnAspThrLysLeuCysLysAspTrpAlaG 291
Db 851 TGGATGATAGTACTCTGCTCTCCCAACATCATGATATGACACTATCCAGAGACTGGAAAG 910
Qy 291 luLysGluProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluS 311
Db 911 ATGAACAGCCCATATACATCTGTATGATACAGCTCTCAGGAAGCTCTCTACCTCAGT 970
Qy 311 exPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValVals 331
Db 971 CTTTCTCATATACACACAGGCTTAGAANAACCTCAAGTCAATGTTGTTGTT 1030
Qy 331 erProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuG 351
Db 1031 CCCCCTCTATATACCTGTTGAAGACTGTCTGCATCAAGGAGATCTCAGCTGGTCTCTCG 1090
Qy 351 luArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArg 371
Db 1091 AGAACATCTCCAAATGAGTCTGATAGATACAAAGTCTTCCATCTCTGATAGAAAATCACC 1150
Qy 371 luLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnL 391
Db 1151 AGCTGTTTGACCAATGCCAGGCCCTCTGTGCTGCTCCCTGTGTGAGGCTCTACAGC 1210
Qy 391 euGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisA 411
Db 1211 TACAGAAGAACTGGGAAGAGATGACCTACCTGCCAGCTCTCACCGGTTTGTATG 1270
Qy 411 laAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuG 431
Db 1271 CCACGTGTGTGTTTCCACAGCTCACCTTGAAAGAGGCTTCCACAGAGCGCTCTCAGTCA 1330
Qy 431 luGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgL 451
Db 1331 AAGAACAATACTCTAGTGGGTTGTGATGATGGCAGCTGAAGAGTGTGGACCATGA 1390
Qy 451 ysSerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgA 471
Db 1391 GGTGGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1450
Qy 471 laLeuPheHisMetAsnIleLeuLeuProAspSerHis---CysGluGluTyrTyrThrP 490
Db 1451 CCCTCTTTCACATGAACATCTCTTCCAGGTTGGCCACACAGTGCAGCAGTGTATGTTT 1510
Qy 490 hePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu- 509
Db 1511 TCTCCACCTCAGCTGAGGATTTCTTGTGCTCTATATATATATATATATATATATATATAT 1570
Qy 510 --GluIleGluProAlaLeuCysProLeuTyrValCulLysThrLysArgSerMetGluL 529
Db 1571 AGGAATGGATCAGCATTTTGC-----TTCATTGAAACCAAGAGGATCATGGAGG 1624
Qy 529 euLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuV 549
Db 1625 TGAAGAGAACTGAC--GACACTCGCTCCTCGGATGAAGCGTTTCTTATTGGCTCTCA 1681
Qy 549 alSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyV 569
Db 1682 TGAACAGGATATCTTGAAGACTCTGGAGGTTCTGTTTGAATATATCCCGTGAATCCAACTG 1741
Qy 569 alLysGlnLysLeuLeuHisIleTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrP 589
Db 1742 TTGACGAGAGCTCCAACTGGGTTCTCTCTGATAGCTCAGCAGGTCAATGGCACCGCC 1801
Qy 589 roGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValA 609
Db 589 roGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValA 609

Db 1802 CAATGGACACCCCTGGATGCCTTCTATTGTCTATTGTAGTCTCAGGATGAAGAGTTTCTTG 1861
Qy 609 rgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuI 629
Db 1862 CGCGGGCTCTCAACGCTTCCAGAGAGTGTGGCTGTGATTTAACAGAGAGTGGACTTGA 1921
Qy 629 leAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValL 649
Db 1922 AGGTCTCTTCTACTGTCTCAAGCAGTGTGAGAACTTGAAGGCAATCCGGGTGGATATCA 1981
Qy 649 ysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMetA 669
Db 1982 GAGACCTCTCTCGTATAGATAATACTCTCAGAGCTGTCCCTGTGTGTACT-----GCC 2035
Qy 669 rgAsp-----LysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuG 686
Db 2036 GGGAGACACAACTGAAGCCCTCTCATGAGTGGTGGGNAACCTCTGCTGTGTGCTTG 2095
Qy 686 lyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaM 706
Db 2096 GCAGCTCCGGAACCTTGAGGAGCTGAGCTTGGCGCAGCAGCATCTCAGTCAACGGGCCA 2155
Qy 706 etLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheA 726
Db 2156 TGAAGATACTGTGCTCGAGCTCGGAATCAGTCTCGCAGAACTACAGAACTGACGTTTA 2215
Qy 726 rgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgA 746
Db 2216 AGAGTGAGAGGTAGTGTCTGGCTGAAACATCTCTGGAAGCTCCTTTTAGCAATCAA 2275
Qy 746 snLeuArgSerLeuAsnLeuGlyIleThrHisLeuLysGluGluAspValArgMetAlaC 766
Db 2276 ACTTAAAGTACTCAATCTAGGGAACACTCCCATGAAGGATGATGACATGAAGTTAGCT 2335
Qy 766 ysGluAlaLeuLysHisProLysCysLeuLeuGlnSerLeuArgLeuAspCysGlyL 786
Db 2336 GCGAAGCGCTGAAACATPCCAAAGTGTCCGTGGAGACTCTGAGGTTGAGTTCTCTGTGAGT 2395
Qy 786 euThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysS 806
Db 2396 TAACCATCATGTTGTTATGAGATGATCTCCAGCTTCTTATTTCACACACGAGGTAAAGT 2455
Qy 806 erLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaL 826
Db 2456 GTCTCAGCTGCTCCCAAAATAGAGTGGAGTAAAAAGCATGATATCCTTTGGGAATGCCT 2515
Qy 826 euArgValSerGlnCysAlaLeuGlnLysLeuLeuGluAspCysGlyIleThrAlaT 846
Db 2516 TGAGTAGCTCAATGTGTCTACTGCAAAAGTTGATCTGGAACACTGTGGCTTCACACCTG 2575
Qy 846 hrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysL 866
Db 2576 CCAGCTGCCACCTTCTGCTCTCAGCCCTTTTCAGCAACCAAGACTTGACACCTGTGCC 2635
Qy 866 euSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuP 886
Db 2636 TGTCAACACAGCTCTGGGACTGAAGAGTGCACAGACTGTGTCTCAGTCTCTGAGGAATC 2695
Qy 886 roHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysG 906
Db 2696 CAGAAATGTGCTCTCAGCGCTGATCTACTGAATCACTGCAACACTTGTAGATGATGCTTATG 2755
Qy 906 lyPheLeuAlaLeuAlaLeuMetClyAsnSerTrpLeuThrHisLeuSerLeuSerMetA 926
Db 2756 GTTCTCTGGCAATGAGACTTGCACCAACCAAGCTGACCCACTGAGCTGAGCTGACCATGA 2815
Qy 926 snProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysH 946
Db 2816 ACCCGTAGGGATGGTCAATGAGCTACTGTGTGAAGCTTTTAAAGAACCTACTTGT 2875
Qy 946 isLeuGlnAspLeuGluValLysCysHisLeuThrAlaAlaCysCysGluSerLeuS 966
Db 2876 ACCTTCAAGAACTGGAACCTAGTGGACTGCCAACTCACAACAGAACTGCTGGCAGGACCTGG 2935

```
QY 966 exCysValIleSerArgSerHISLeuLysSerLeuAspLeuThrAspAsnAlaLeuG 986
Db 2936 CCTGTATGATACACACACACAGCACTTAAAGATTGGATCTTGGTAAACACCCCTGG 2995
QY 986 lyAspGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrA 1006
Db 2996 GTGACAAAGGAGTATACACCTGTGTGAGGAGCTGAACAAAGTAGCAGCTCCCTGAGGA 3055
QY 1006 rgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaL 1026
Db 3056 GACTTGGGTGGGGGCATGTAAAGTTGACTTCCAAATGCTGTGAGGCACTGTGCATGGCCA 3115
QY 1026 euSerCysAsnArgHISLeuThrSerLeuAsnLeuValGlnAsnAspPheSerProLysG 1046
Db 3116 TCTCTTGGAAACCCCTCAGCTGACAGCCTAAACCTGGTGGAGAAATGACTTCAATCATCGG 3175
QY 1046 lyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleleGlyL 1066
Db 3176 GGATGTTGAAGCTGTGCTCGTTCCTCAATCCCTGTCTCTAACTGGGATATATTGCC 3235
QY 1066 euTrpLysTrpGlnTrpProValGlnIleArgLysLeuLysGluGluValGlnLeuL 1086
Db 3236 TGTGGAAGCAGGAGTACTATCCCGAGTGAAGACAGCTGGAGAAAGTTGAGTTGTCA 3295
QY 1086 ysProArgValIleAspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3296 AGCCCCAGCTGTGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3351

RESULT 2
LOCUS BC021272 3359 bp mRNA linear HTC 17-DEC-2003
DEFINITION Homo sapiens cDNA clone IMAGE:4811303, containing frame-shift errors.
ACCESSION BC021272
VERSION BC021272.2 GI:33878145
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3359)
Strausberg,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butcherfield,X.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3359)
Strausberg,R.
Direct Submission
Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK
COMMENT On Aug 19, 2003 this sequence version replaced gi:18204229.
```

Contact: MGC help desk
Email: gcaps-x@mail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
AKhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granillo,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,K., Stantripop,S., Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: g Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
1..3359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4811303"
/issue_type="Placenta, choriocarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Alignment Scores:
Pred. No.: 3,359-133 Length: 3359
Score: 1422.00 Matches: 349
Percent Similarity: 48.00% Conservative: 204
Best Local Similarity: 30.30% Mismatches: 448
Query Match: 20.10% Indels: 151
DB: 11 Gaps: 15
US-10-066-521-6 (1-1344) x BC021272 (1-3359)
QY 9 PheSerSerTrpGlyLeuGlnTrpCysLeuTrpGluLeuAspLysGluGluPheGlnThr 28
Db 76 TTCTCTGATTGTTGGTCTTATGTTGTTATCTGGAGGAGCTCAAAAAGGAGGATTCAGGAAA 135
QY 29 PheLysGluLeuLeuLysLysLysSerSerGluSerThrThrCysSerIleProGlnPhe 48
Db 136 TTTAAAGAACATCTCAAGCAAAATGACTTTGCAGCTTGAACCTCAAGCAGATTCCTGGACT 195
QY 49 GluIle-GluAsnAlaAsnValGluCysLeuAlaLeuLeuLeuHISGluTrpTyrGlyAl 68
Db 196 GAGGTCAAAAAGACATCCCGGAGAACTTGCACACCTCTTGCATCAAGCAGCTATGAGA 255
QY 68 aSerLeuAlaTrpAlaThrSerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSe 88
Db 256 ACNACAGCTTGGACATCACTTAAAGATCTTTCAAAAGATG----- 298
QY 88 rGluLysAlaArgAspAspMethLysLysIleSerGlnAlaMetGluGlnGluGlyAlaTh 108
Db 299 -----GATAGAAAGGATCTCTGCATGAAGGTCAATGAGGAGAGACAGC 342
QY 108 rAlaAlaGluThrGluGluGlnGluSerGlnAlaMetGluGlnGluGlyAlaThrAl 128
Db 343 ATACACAAAGACCTATCAAGCTCAACGCAAGCAGAAATTCAGCCGC----- 388
QY 128 aAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMe 148

Db 389 -----TTATGGTCCACCAAGTCTGTCACTGA 414
Qy 148 tThrLysPheAlaGluGluAAspValArgSerPheGluAenThrAlaAlaAspTr 168
Db 415 GATTCACTTATCTTTGAGGAGAAAGTCAAGCAAGAA----- 451
Qy 168 pProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProAr 188
Db 452 ----GAATGTGACCATTTGGACCGCTTTTGTCTCCCAAGGAAACTGGGAAACAGCCACG 507
Qy 188 gThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaAaGArgIleVa 208
Db 508 TACAGTGATTATCAAGGACCAAGAAATGGAAAAACAGACACTCTCTGATGAAGCTGAT 567
Qy 208 lleuCysTrpAlaGlnGlyGlyLeuTyrglnGlyMetPheSerTyrrValPhePheLeuPr 228
Db 568 CATGCGCTGGTCGGACAACAAGATCTTCGGGATAGGTTCTGTACACGTTCTATTTCTG 627
Qy 228 oValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArgGluTr 248
Db 628 CTGCAGAGAACTGAGGGAGTTGCCGCCAACGAGTTTGGCTGACTTGTATTCAGAGAGTG 687
Qy 248 pProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIl 268
Db 688 GCCTGACCCCGCTGCTCTATACAGAGATGCTGTCTCAACCGGAGACACTTGTGTCGT 747
Qy 268 eileAspGlyPheAspAspLeuGlySerValLeuAenAen---AspThrLysLeuCyL 287
Db 748 CATCGACAGCTTCGAAGAGCTGCAGCGCGGCTTGAACGAAACCCGATTCCGATCTGTGTGG 807
Qy 287 shAspTrpAlaGlyLysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLe 307
Db 808 TGACTTGTATGGAGAAACCGCGGTGCAGGTGCTTCTGAGCAGTTTGTCTGAGGAAGAAGAT 867
Qy 307 uLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSe 327
Db 868 GTCCCGGAGGCTCCCTGCTCATCGCCATCAAAACCGTGTGCCGAGAGAGCTCCGGGA 927
Qy 327 rGluValValSerProArgTyrrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHi 347
Db 928 TCAGGTGACGATCTCAGAAATCTACACGCGCGGGATTCAACGAGAGATGATAGGTAGT 987
Qy 347 sLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMe 367
Db 988 GTATTTCTGCTGTTTCTTCAAGACCCGAAAGACCCATGGAAGCCTTCAATCTTTGAAG 1047
Qy 367 tAsnAenArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCyEva 387
Db 1048 AGAAAGTGAAACGCTGTTTCCATATGCAAAATCCCGTCTCTCTGCTGGATCTGTGTGAC 1107
Qy 387 lAlaLeuGlnLeuGlnAspValValGlyLysSerValAlaProPheAenGlnThrLeuTh 407
Db 1108 CAGTCTGAAGCAAGAGATGCAGAAAGGAAAGACCTGGCCCTGCCTCCAGAGACTAC 1167
Qy 407 rGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCy 427
Db 1168 CTCTGTGTACTCTCTTCTGCTTTAACTGTTCACACTGAGGGTGCAGGGGCCCGAC 1227
Qy 427 sLeuAenLeuGluArgValValLeuLysArgPheCysArgMetAlaValGluGlyVa 447
Db 1228 TCCGCAAAACCCAGCACCA-----CTGAAGGCCCTGTGCTCCCTGCGAGAGGGTAT 1281
Qy 447 lTrpAenArgLysSerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluse 467
Db 1282 GTGGACAGACATTTGAGTTTGTGAAGACGACCTCCGGAGAAATGGGGTTGTGAGCG 1341
Qy 467 rGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTy 487
Db 1342 TGACATCCCTGGCTGCTGGGCAACCAAGATACTTCTGAAGTACGGGGCGGTGAGAGCTC 1401
Qy 487 rTyrrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrrValLeuGl 507
Db 1402 CTACGTGTTCTCCACGTTGTATCCAGGAGTTCTGTGCCGCTTGTCTATTTGCTCAA 1461

Qy 507 uGlyLeuGluIleGluPro-----AlaLeuCysProLeuTy 519
Db 1462 GAGCACCCTTGATCATCTCACCAGCTGTGAGATGTGTACAGGAATTGTAGTTGCCAA 1521
Qy 519 rValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLe 539
Db 1522 TTTTGAAGAAAGCAGGAGAGCA-----CATTGAT 1551
Qy 539 uTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVa 559
Db 1552 TTTTGGGGTGTCTTAACTGGCCTTTTAAATAAAGAAAGCAAGAAACTTGATGC 1611
Qy 559 lLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpValSerLe 579
Db 1612 GTTTTGTGGCTTCCAACTGTCCCAAGAGATAAAGCAGCAAAATTCACCGACTGCTGAAGAG 1671
Qy 579 uLeuGlyGlnGlnProAenAlaThrProGlyAspThrLeuAspAlaPheHisCyLe 599
Db 1672 CTTAGGGGAGCGTGGCAATCTCAGGACAGGTGGATTCTTGGCGATATTTTACTGTCT 1731
Qy 599 uPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAenSerPheGlnGluValTr 619
Db 1732 CTTTGAATGCAGGATCTGCTTGTGAAGCAGGAGTGAACCTCTCCAGAGAGCTAA 1791
Qy 619 pLeuProIleAenGlnAenLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPr 639
Db 1792 CTTTCATATTATTGACAACGTGGACTTGGTGTCTTCTGCTACTGCTTAAATACTGCTC 1851
Qy 639 oTyrrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGl 659
Db 1852 CAGCTTGAGAAACTCTGTTTTCCGTTCAAAATGTCTTTAAGAAAGAGGATGAACACAG 1911
Qy 659 uAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGl 679
Db 1912 CTCT-----ACGTGGATTACGCTCATC-----TGTGGCA 1944
Qy 679 uAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSe 699
Db 1945 TCACATCTGCTGTGCTCACCAGCGGACCTCAGAGAGCTCCAGGTCAGGAGCAG 2004
Qy 699 rIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLy 719
Db 2005 CACCTCAGGAGTCGACCTTGTGACCTGCTGTGAACCTGAGCTGAGGATCCACGCTGCG 2064
Qy 719 sIleGlnThrLeuMetPheArgAenAlaGlnIleThrProGlyValGlnHisLeuTrpAr 739
Db 2065 CTTTCAAGCTTGAATAATAACGTTTCTTCTGGCCAGAGTGTCTGCTCTTTGA 2124
Qy 739 gIleValMetAlaAenArgAenLeuArgSerLeuAenLeuGlyThrHisLysLysGl 759
Db 2125 GGTGCTCTTTTATCAGCCAGACTTGAATAACCTGAGCTTCAACCTCAGAAACTCTCTCG 2184
Qy 759 uGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLe 779
Db 2185 TGATGATCAGTCCCTCTGTGATGCTTGAATCCCA----- 2224
Qy 779 uArgLeuAspCysCysGlyLeuThrHisAlaCysTyrrLeuLysIleSerGlnIleLeuTh 799
Db 2224 ----- 2224
Qy 799 rThrSerProSerLeuLysSerLeuLeuAlaGlyAenLysValThrAspGlnGlyVa 819
Db 2225 -----GCAGGCAAC----- 2233
Qy 819 lMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGl 839
Db 2234 -----GTCAAGAGAGCTAGCGCTGT 2253
Qy 839 uAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerLeuAlaLeuValSerAenAr 859
Db 2254 AAATTGTACCTCTCACCCTTGAATTGTAAGTCTTGTGCGCTTCTTAACCAACAACAA 2313

```
Qy 859 gSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValaenLeuLe 879
Db 2314 GAAGCTGACGATCTGATGATCTCTGCAACAGTTA---GACACAGCGGTGCCCTTTT 2370
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAenGlnCysHi 899
Db 2371 GTGTGAAGCCCTGTGCAGCCAGACAGCGTCTGCTATACCTGATGATGTTGGCTTTCTGCCA 2430
Qy 899 sLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuTh 919
Db 2431 CCTCAGCAGCAGCTGCTCGCAATACATCTCTGAATGCTTCGTGAACAGACGCGTGC 2490
Qy 919 rHisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallylsLeuLeuCysGluVa 939
Db 2491 CTATCTAGACCTCAGTGCATGCTCCTGAAGAGCAGAGGACTGAAACTCTCTCGGAGGC 2550
Qy 939 lMetArgGluProSerCysHisLeuGlnAspLeuGluLeuVallylsCysHisLeuThrAl 959
Db 2551 CTTGAAACATCCGAGCTGCTCGCTGGATTCACTGTGTTGGTAAATGTTTATCACTGC 2610
Qy 959 aAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuIlyssSerLeuAs 979
Db 2611 TCGTGGCTCGAAGACCTTCGCTCTGCTCATCAGCAATCAAAACCTGAAGATTCTGCA 2670
Qy 979 pLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGl 999
Db 2671 AATGGGTGCATGAAATCGAGATGCGGTGTCAGCTGTGTGTCGGCTCTGACGCA 2730
Qy 999 nLysAsnSerValLeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysCy 1019
Db 2731 TACGGATTGCCGCTTAGAGATTCTGGGTGGAAGATTGGGTAAACGAGCACCTGCTG 2790
Qy 1019 sGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGl 1039
Db 2791 TAAGGATCTCGCGTCTGTCTCACTCGAGTAAGACCCCTCGAGCAGCTCAACCTGACCTT 2850
Qy 1039 nAsnAsnPheSerProLysGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSe 1059
Db 2851 GAACACCTTGGACACACAGGGGTGGTTGTACTCTGTGAGCCCTGAGACACCCAGAGTG 2910
Qy 1059 rAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLe 1079
Db 2911 TGCCCTGCGAGTGTGCGGCTGAGAAAACCTGATTGTGAGGAAACCCAGGACACTTCT 2970
Qy 1079 uGluGluValGlnLeuLeuLysProArgValValIleAsp----- 1092
Db 2971 GACGGCTGAGGAGAGAGAAATCCTAACCTGACCATCAGATGACTGTGACACAATCAC 3030
Qy 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLy 1105
Db 3031 AAGGTAGAGATCTGATTGCCAGAACCTGG----- 3061
Qy 1105 sIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGl 1125
Db 3062 ---GCTCTGACTCGAACACCTGCAAGAGCAGGAGCTGGGAC---CGTTACTTTACATGA 3114
Qy 1125 yMetAsnProGluGlnIlyssLysArgValSerLeu 1136
Db 3115 CACTGCACCCAGGAGATACAAATCATTTGACACTC 3148

RESULT 3
LOCUS BC012789 3360 bp mRNA linear HTC 17-DEC-2003
DEFINITION Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift errors.
ACCESSION BC012789
VERSION BC012789.2 GI:33872518
KEYWORDS HTC
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3360)
```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zdobych, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3360)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:15215377.

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ARCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: anadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettenan, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 25 Row: k Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 19745161

This clone has the following problem: frame shifted.

Location/Qualifiers

1..3360

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3957172"

/tissue_type="placenta, choriocarcinoma"

/clone_lib="NIH MGC 21"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

Qy	29	PhelysGluLeuLeuLeuLysLysSerSerGluSerThrThrCysSerIleProGlnPhe	48
Db	136	TTTAAAGAACATCTCAAGCAAAATGACTTTGGAGCTTGAATCAAGCAGATTCCTCGGACT	195
Qy	49	GluIle-GluAsnAlaIenValGluCysLeuAlaLeuLeuLeuHisGluIuTyTyGlyAl	68
Db	196	GAGGTCAAAAGAGCATCCCGGAGAACTTCGAAACCTCTTGATCAAGCACTATGAAGA	255
Qy	68	aseLeuAlaTrpAlaThrSerIlePheGluAsnMetAsnLeuArgThrLeuSe	88
Db	256	ACAACAAGCTTGGAACTAATCTTTCAAGATG-----	298
Qy	88	rGluLysAlaArgAspMetLysLysIleSerGlnAlaMetGluGlnGluGlyAlaTh	108
Db	299	-----GATGAAAGGATCTCTGATGAAGTCTATGAGGAGAGAACAGG	342
Qy	108	rAlaAlaGluThrGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAlaThAl	128
Db	343	ATACACAAGACCTATCAAGCTCAGCAACAGCAAAATTCAGCGC-----	388
Qy	128	aAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyTySerHisValMe	148
Db	389	-----TTATGGTCCAGCAAGTCTGTCACTGA	414
Qy	148	tThrLysPheAlaGluGluAspValArgSerPheGluAsnThrAlaAlaAspTr	168
Db	415	GATTCACCTATACTTTGAGGAGGAAGTCAAGCAAGAA-----	451
Qy	168	pProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProAr	188
Db	452	---GAATGTGACCATTTGGACCGCTTTTGCTCCCAAGAAACTGGGAAACAGCCACG	507
Qy	188	gThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleVa	208
Db	508	TACAGTGATTATTCAAGACCACCAAGAAATTTGGAAAAACGACACTCTCTGATGAAGCTGAT	567
Qy	208	lLeuCyTrpAlaGlnGlyLeuTyGlnGlyMetPheSerTyTyValPhePheLeuPr	228
Db	568	GATGGCTTGGTCGGAACAACAGATCTTTTCGGGATGAGTTCTGTACACGTTCTATATTCGT	627
Qy	228	oValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArgGluTr	248
Db	628	CTGCAGAGAACTGAGGGAGTTGCCGCCAACAGAGTTGGCTGACTGATTCACAGAGATG	687
Qy	248	pProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIl	268
Db	688	GCCTGACCCCGCTCTCTATAACAGAGATCGTGTCTCAACCGGAGAGACTCTTGTTCGT	747
Qy	268	eIleAspGlyPheAspAspLeuGlySerValLeuAsnAsn---AspThrLysLeuCyLe	287
Db	748	CATCGACAGCTTCAGAGACTGCAGGGCGGCTTGAGCAACCCGATTCGGATCTGTGTGG	807
Qy	287	sAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLe	307
Db	808	TGACTTGATGAGAAACGGCCGGTGCAGGTCTTCTGAGCAGTTTGTCTGAGAAAGAGAT	867
Qy	307	uLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSe	327
Db	868	GCTCCCGGAGGCCCTCCCTGTCTCATCGCATCAAAACCCGTGTGCCGAGAGGACTCCGGGA	927
Qy	327	rGluValValSerProArgTyTyLeuLeuValArgGlyIleSerGlyGluGlnArgGlieHi	347
Db	928	TCAGGTGACGATCTCAGAAATCTACCAAGCCCCGGGGATTCAACAGAGTGTAGTGTAGT	987
Qy	347	sLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMe	367
Db	988	GTATTTCTGCTGTTTCTTCAAAGACCCGAAAAAGACCATGGAAGCTTCAATCTGTGAAG	1047
Qy	367	tAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysVa	387
Db	1048	AGAAAGTGAACAGCTGTTTTCATATGCAAAATCCGCTCTCTGCTGGATCTTGTGTGAC	1107

Qy	387	lAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuTh	407
Db	1108	CAGTCTGAAGCAAGAGATGCAGAAAGAAAGACCTGGCCCTGACCTGCCAGGACCTAC	1167
Qy	407	xGlyLeuHiSaLaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCy	427
Db	1168	CTCTGTGTACTCTCTTTCGTCTTAACCTGTTCACACCTTGAGGGTCCCGAGGCCGAC	1227
Qy	427	sLeuAsnLeuGluGluArgValValLeuLeuysArgPheCysArgMetAlaValGluGlyVa	447
Db	1228	TCCGCAAAACCAGCACACAG-----CTGAGGGCCCTGTGTCTCCCTGCTCGAGAGGGTAT	1281
Qy	447	lTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSe	467
Db	1282	GTGGACAGACACATTTGAGTCTTTGTGAAGACGACCTCCGAGAGAAATGGCGTGTGTGACGC	1341
Qy	467	xGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTr	487
Db	1342	TGACATCCCTGGCGTGTCTGGCCACCAAGATACTTCTGAAGTACGGGAGCGTGAGAGCTC	1401
Qy	487	rTyrThrPhePheHisSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGlu	507
Db	1402	CTAGTGTCTCTCCACGTGTGTATCCAGGAGTCTGTGCCGCTGTGTCTATTGTCTCAA	1461
Qy	507	uGlyLeuGluIleGluPro-----AlaLeuCysProLeuTy	519
Db	1462	GAGCCACTTGTATCATCTCCACGCTGTGAGATGTGTACAGAAATTGTAGTTGCCAA	1521
Qy	519	rValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLe	539
Db	1522	TTTTGAAAAAGCAAGGAGACA-----CATTCGGAT	1551
Qy	539	uTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVa	559
Db	1552	TTTTTTGGGTGTCTTAAGTCTTAAATAAATAAAGCAACAGAAAACTGGATGC	1611
Qy	559	lLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTrpValSerLe	579
Db	1612	GTTTTTTGGCTTCCAACTGTCCCAAGAGATAAAGCAGCAAAATTTCCACAGTGCCTGAAGAG	1671
Qy	579	uLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLe	599
Db	1672	CTTAGGGGCGGTGGCAATCCTCAGGACAGGTGGATTCCTTGGCGATATTTTACTGTCT	1731
Qy	599	uPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTr	619
Db	1732	CTTTGAAATGCAGATCCTCCCTTTGTGAAGCAGGCAGTGAACCTCCTCCAGAGAGCTAA	1791
Qy	619	pLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPr	639
Db	1792	CTTTCATATTATGACAACTGACCTGGTGTGGTGTCTGCCTACTGCTTTAAATATCTGCTC	1851
Qy	639	oTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGl	659
Db	1852	CAGCTTGAGAAACTCTGTTTTTCCGTTCAAAATGTCTTTTAAGAAAGAGATGAACACAG	1911
Qy	659	uAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGl	679
Db	1912	CTCT-----ACGTCGGATTACAGCTCATC-----TGTGGCA	1944
Qy	679	uAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSe	699
Db	1945	TCACATCTGCTGTGTCTACCCAGCAGCGGCACCTCAGAGAGCTCCAGTGCAGGACAG	2004
Qy	699	rIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLy	719
Db	2005	CACCTCAGCGAGTCGACCTTTGTGACCTGGTGTAACTACAGCTCAGGATCCCAAGCTGTGC	2064
Qy	719	sIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpAr	739
Db	2065	CCCTTCAGAGCTTGGAAATAAATACGTTTCTCTTTCTGGCCAGAGTTCCTGCTCTTTGA	2124
Qy	739	gIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGl	759

```
Db 2125 GGTGCTCTTTTATCAGCCAGACTTGAATACCTGAGCTTACCTCCCTCAGAAACTCTCTCG 2184
Qy 759 uGluaspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLe 779
Db 2185 TGATGACATCAGGTGCTCTGTGATGCTTGAATACCA 2224
Qy 779 uArgLeuaspCysCysGlyLeuThrHisAlaCysTyrlsLeuLysIleSerGlnIleLeuThr 799
Db 2224 ----- 2224
Qy 799 rThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVa 819
Db 2225 -----GCAGGCAC----- 2233
Qy 819 lMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGl 839
Db 2234 -----GTCAAAGAGCTAGCGCTGGT 2253
Qy 839 uAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnAr 859
Db 2254 AAATGTCACTCTCACCCATGATGTGAAGTCTTGTGCGCTTCTTAACCAACAA 2313
Qy 859 gSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLe 879
Db 2314 GAAGTGAAGTATCTGATGATCTCTGCAACAGTTA---GACACAGCGTGCCCTTTT 2370
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHi 899
Db 2371 GTGTGAAGCCCTGTGCAGCCAGACAGCGTCTGTGTATCTGATGTTGGCTTTCTGCCA 2430
Qy 899 sLeuAspThrAlaGlyCysGlyPheLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2431 CCTCAGCAGCAGTGCTCGAATACATCTCTGAAATGCTTCTGCGTAAACAGACGCTGCG 2490
Qy 919 rHisLeuSerLeuMetAsnProValGluAsnGlyValLysLeuLeuCysGluVa 939
Db 2491 CTATCTAGACCTCAGTGCCATGCTCTGAGAGCAGAGGACTGAAACTCTCTCGAGGC 2550
Qy 939 lMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAl 959
Db 2551 CTTGAAACATCCGAGCTGCTCGCTGATCTCACTGTGTTGGTAAATGTTTATCACTGC 2610
Qy 959 aAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAs 979
Db 2611 TGCTGGCTGTGAAGACCTCGCTCTCTCATCAGCAATCAAACCTGAAGATCTGCA 2670
Qy 979 pLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGl 999
Db 2671 AATTGGGTGCATGAAATCGAGATGCGGTGTGCAGCTGTGTGTCGGGCTCTGACGCA 2730
Qy 999 nLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCy 1019
Db 2731 TACGGATTGCGGCTTAGAGATTCTTGGGTGGAAGAATGTGGGTAAACGAGCACCTGCTG 2790
Qy 1019 sGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGl 1039
Db 2791 TAAGATCTCCGGCTGTGTTCTACCTGATGATGAACCTCGAGCAGCTCAACCTGACCTT 2850
Qy 1039 nAsnAsnPheSerProLysGlyMetLysLeuCysSerAlaPheAlaCysProThrSe 1059
Db 2851 GAACACCTTGGACACACAGGGGTGTTGTACTCTGTGAGGCCCTGAGACACCCAGAGTG 2910
Qy 1059 rAsnLeuGlnIlelleGlyLeuTrpLysTrpGlnTyrlProValGlnIleArgLysLeuLe 1079
Db 2911 TGCCCTGAGGTGCTCGGGCTGAGAAAACCTGATTTTATGAGGAAACCCAGGCACTTCT 2970
Qy 1079 uGluGluValGlnLeuLysProArgValValIleasp----- 1092
Db 2971 GACGCTGAGGAGAGAGAAATCTTAACCTGACCATACAGATGCTGTGACACAATCAC 3030
Qy 1093 -----GlySerTrpHisSerPheAspGluAspArgHisLy 1105
```

```
Db 3031 AAGGGTAGAGATCTGATTGCGAGGAACCTGG----- 3061
Qy 1105 sIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCysAlaLeuLeuTrpGl 1125
Db 3062 ---GCTCTGACTCGAACACCTGCAAGGACAGGACTGGGAC---CGTTACTTACATGA 3114
Qy 1125 yMetAsnProGluGlnLysLysArgValSerLeu 1136
Db 3115 CACTGCACCCAGAGATACAAATCATTTGACATC 3148

RESULT 4
CK000513 791 bp mRNA linear EST 26-NOV-2003
AGENCOURT_16368905 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30708637 5', mRNA sequence.
CK000513
CK000513.1 GI:38526547
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
REFERENCE
1 (bases 1 to 791)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: dsgerb@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAMI075 row: c column: 14
High quality sequence stop: 688.

FEATURES
source
1..791
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708637"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/vector="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5KB. Adaptors 5' (AATTCGACGAGG)3' and 5' d
(CCTGTGCGG)3'. 3' linker sequence - GCGGCGCTGAGAGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'd
(TAATACCTCACTAAGGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

ALIGNMENT SCORES
Pred. No.: 2,2e-113 Length: 791
Score: 1217.00 Matches: 255
Percent Similarity: 98.08% Conservative: 1
Best Local Similarity: 97.70% Mismatches: 5
Query Match: 17.20% Indels: 5
DB: 14 Gaps: 0

ORIGIN
```


US-10-066-521-6 (1-1344) x CK000513 (1-791)		High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636	
TITLE	JOURNAL	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	
REFERENCE	MEDLINE	20499374	
AUTHORS	PUBMED	11042159	
TITLE	JOURNAL	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Maeumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)	
REFERENCE	MEDLINE	20530913	
AUTHORS	PUBMED	11076861	
TITLE	JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
REFERENCE	MEDLINE	20530913	
AUTHORS	PUBMED	11076861	
TITLE	JOURNAL	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	
REFERENCE	MEDLINE	20530913	
AUTHORS	PUBMED	11076861	
TITLE	JOURNAL	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takanashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT	JOURNAL	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. 3577 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:E330028A19" /db_xref="MG:2426651" /db_xref="taxon:10090"	
FEATURES	source		

[illegible]

QY 384 LeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsn 403
D 1230 ATGGTGGCTACTGTCTAAAGAGAGAGATAGAGAAGGGAAGAGACCCAGCTCCATTGCG 1289
QY 404 GlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyVal 423
D 1290 CGTTGACTACCTCCCTGTATACCACTCACATCTTCAATTTGTTTCATTCCTCCCAAAATGCC 1349
QY 424 ValArgArgCysLeuAsnLeuGluGluArgValValLeuLeuYsArgPheCysArgMetAla 443
D 1350 CATTCTCCAAGTAAGAAAGCCAGACCCAG-----CTGCAGGGCTTAATGTTCTCTGCCC 1403
QY 444 ValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGly 463
D 1404 GCTGAGGGCATGTGACTGACACAGCTTTGTGTTGTGAGGAGGCTCTCAGGAGAAATGGG 1463
QY 464 LeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHis 483
D 1464 ATCATGGACTCTGCATCCCCATCTACTGGACATTTGGAAATGCTTATAAATATATCAGGGAA 1523
QY 484 CysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyr 503
D 1524 TCTGAAAATCTTATATATTTCTCATCTCACTGTTTCAGGAGTCTGTGAGCTATCTTT 1583
QY 504 TyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThr 523
D 1584 TATCTGCTAAAGAGCCACGTGACACCCCTAGCCAGGAGGTTAAAGATATAGAGAAATCTC 1643
QY 524 LysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArg 543
D 1644 ATGTTTGCATTT---CTAAGAAAGTC-----AAAGTACAGTGGATTTCTTTGGCTCT 1694
QY 544 PheLeuPheGlyLeuValSerGluAspValArgProLeuGluValLeuLeuGlyCys 563
D 1695 TTCATCTTTGGCTTTTACATGAATCAGAACAAAAAGCTAGAGCGCTTTTTTTGGCCA 1754
QY 564 ProVal-ProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnG 583
D 1755 CCAGTTGTCCTCCCAAGAAATAAACGTCAGTTGTATCAGTCGCTGGAACTATAAGTGGCA 1814
QY 583 nProAsnAlaThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrG 603
D 1815 CGAAGAGCTTCAAGAACACATAGATAGGCTGATGAAGCTGTTTACTGCTGTGTGAGATGA 1874
QY 603 nAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAs 623
D 1875 CGATGACACCTTCTTAGTAGAAGCTATGAACCTGTATGGAACAGATTAACCTTTGTGGCTAA 1934
QY 623 nGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgL 643
D 1935 GGATTTATCTGATGTTATTTGTTGCTGCCACCTGCTTAAACACACTGTTTACACTGAAGAA 1994
QY 643 sIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProVa 663
D 1995 ACTATCTCTTTTCAACCCCAAGATGCTCTGAGT----- 2025
QY 663 lValProLeuTrpMetArgAspLysThrLeuIleGluGlu-----GlnTrpGluAs 680
D 2026 -----GGAGCTCAAGAACACACGCTATATGAAAGGCTGCTCACTGCTTGAATCA 2075
QY 680 pPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuLysSerI 700
D 2076 TATTTGCTCTGTTTCATATCAGTAAGACACTTCAGGACTCCGAATGAAGACACTAA 2135
QY 700 eLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysI 720
D 2136 TCTCAGTGAATCAGCCTTTTTCAGCTTGTATATAATTAATTAAGTACCACCACTACCCCT 2195
QY 720 eGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgI 740
D 2196 TAAAGCTTCTGTGGCAAAATATGTGTTCTTCGTA---TGTGAAAGATGACTGTTCTTGA 2252

QY 740 eValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluG 760
D 2253 GTTGATTCAGAACTGCAATTTACAGCACTTTGAACCTCAGCCTCAATTTCTCTCCACAG 2312
QY 760 uAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuAr 780
D 2313 TGATGTGAACACTGTTGTGTGATGCTGTGAGCCAGGAGTGCACATAGAGAATCTGT 2372
QY 780 gLeuAspCysCysGlyLeuLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 800
D 2373 GGTAGCAGCTTGTAGTATTTTCATCTGATGACTGCAAGGTGTTGCTCTGTCTGATCAG 2432
QY 800 xSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyValMe 820
D 2433 CAACAGACCTTGAAGCATCTTAATTTGGCATCCAAACACCTTG---GACAAAGGAATAGC 2489
QY 820 tProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAs 840
D 2490 CTCCTGTGCAAGTCTTTGTGGCCACCAGACTGCAATTTCTGAGCACTTAGTTTGGCCAA 2549
QY 840 pCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSe 860
D 2550 CTGCTCCCTCAATGAGAATGCTGGGATTACCTTTCAGAAGTCTTAGCGGGAACAAAC 2609
QY 860 rLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCy 880
D 2610 CTTGAGCCACTAGACATCAGCTCCAATGACCTCAAGGATGAAGGCTGAAGGTTCTCTG 2669
QY 880 sArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLe 900
D 2670 TAGGGCTCTGACTCTCCACAGACAGTGTCTGAAAGTCACTAAGTTTGAGACATTTCTCAT 2729
QY 900 uAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHi 920
D 2730 CACCATTTAGTGTGCTCCAGGACCTGGCTGAAGTCTTGAGGACCAACAGCACTGGTCAG 2789
QY 920 sLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluValMe 940
D 2790 CCTACAAGTTTTCGAACAATAAACTAGAAACACACTGGTGTGAAGCTGCTGTGTGATGTAT 2849
QY 940 tArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAl 960
D 2850 AAAACATCCCAACTGCCACTTAGAGGATCTTGGGTTCGAAAGCCTGTGAGCTTAACGTGTC 2909
QY 960 aCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLe 980
D 2910 CTGTTGTGAGGACCTTCATCTACATTTACTCAGTGTAGACTCTGTGGGCAATGAACCT 2969
QY 980 uThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGlnLy 1000
D 2970 ACTGAAGAATGCTTGGACTACAATGGTGTGTTGCTCTGTGTGAAGCTTTTGAACACAGA 3029
QY 1000 sAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysG 1020
D 3030 AACATGTGCCACATPACCTACTTGGATTTGCAAAATTTACTGACTTTTGATCTGAA---ACCCA 3086
QY 1020 uAlaLeuSerLeuAlaLeuSer-----CysAsnArgHisLeuThrSerLeu 1035
D 3087 GGCATTTCTGTTGCTCGAACAGAGAAAATCCATGCTTGGCCATCTTAGCAGCTCTG 3144
RESULT 7
AK016782 1913 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION library, clone:493341P06 product:maternal effect gene, full insert
sequence.
ACCESSION AK016782
VERSION AK016782.1 GI:12855702
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

Db      833  ----- 833
Qy      824 AspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIle 843
Db      833  ----- 833
Qy      844 ThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHis 863
Db      834  ----- 834
Qy      864 LeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMet 883
Db      895 CTGTGCTGCTGTAACACACAGCTGGGGCTGAGAGAGTGCAACCTGTGTGCTGCTGCTG 944
Qy      884 ArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAla 903
Db      945 AGGAATCCAGATGTGCTCTCCAGCGCTGATGTAATCATTTGCAACATTGTAGATCAT 1004
Qy      904 GlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeu 923
Db      1005 GCTTATGGCTTCTGCGCATTTGAGACTTGTCAACACACACAAAGCTGACCCACCTGAGCCTG 1064
Qy      924 SerMetAsnProValGluAsnGlyValLysLeuLeuCysGluValMetArgGluPro 943
Db      1065 ACCATGAACCCGTAAGGAGTGTGCAATGAAGTACTGTGTGAAGCTTTTAAAGGAACCT 1124
Qy      944 SerCysHisLeuGlnAspLeuGluValLysCysHisLeuThrAlaAlaCysCysGlu 963
Db      1125 ACTTGTACCTTAAAGAACTGAACTAGTGACTGCGCACTTATGCAGAACTGCTGTGAG 1184
Qy      964 SerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsn 983
Db      1185 GACCTGGCTGTATGATCACAAACCAAGCACTTAAAGAGTTTGGATCTTGGTAACAAC 1244
Qy      984 AlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerVal 1003
Db      1245 GCCCTGGGTGCAAGAGGAGTATACCTGTGCGAGGAGCTGAGCAAAAGTAGAGCTCC 1304
Qy      1004 LeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSer 1023
Db      1305 CTGAGGAGACTTGGTTGGGGCATGTGATGTGACTTCCATTGCTGTGAGTCATTGTCA 1364
Qy      1024 LeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSer 1043
Db      1365 TTGGCCATCTCTTGAACCCCTCACCTGAACAGCTTAAACCTGATGATGATGACTTCAGT 1424
Qy      1044 ProlLysGlyMetMetLeuLysLeuSerAlaPheAlaCysProThrSerAsnLeuGlnIle 1063
Db      1425 ACATCGGGGAAGTGAAGCTGTGCTGCGTTCCAAATGCCCTGCTCTTAACCTGGGGATA 1484
Qy      1064 IleGlyLeuTrpLysTrpGlnTrpProVal-----GlnIleArgLysLeuLeu 1079
Db      1485 ATTGGGTC-TGGCACTTGGGCTACCTGCATTAACTTTCATGCAATTAAAGAACTA 1543
Qy      1080 GluGluValGlnLeu-----LeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db      1544 AGGACAGAGTATTAAACAATCTGTCTCCAAATTTAATC-----CATACA 1588
Qy      1098 Phe 1098
Db      1589 TTT 1591

RESULT 8
BC036506
LOCUS      3793 bp      mRNA      linear      HTC 19-NOV-2003
DEFINITION Homo sapiens NACHT, leucine rich repeat and PYD containing 11, mRNA
(cDNA clone IMAGE:5265613), with apparent retained intron.
ACCESSION BC036506
VERSION   BC036506.1 GI:23243476
KEYWORDS  HTC.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 3793)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullihy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Viallon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalios, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
PUBMED

22388257
 12477932

REFERENCE
AUTHORS

2 (bases 1 to 3793)
 Direct Submission
 Strausberg, R.
 Submitted (09-AUG-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-@email.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 73 Row: m Column: 9
 This clone has the following problem: retained intron.

FEATURES
source

Location/Qualifiers
 1..3793
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5265613"
 /tissue_type="Testis"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /notes="Vector: pbluescript"

ORIGIN

Alignment Scores:
 Pred. No.: 1-24e-90 Length: 3793
 Score: 1008.00 Matches: 269
 Percent Similarity: 49.47% Conservative: 200
 Best Local Similarity: 28.38% Mismatches: 402
 Query Match: 14.25% Indels: 77
 DB: 11 Gaps: 19
 US-10-066-521-6 (1-1344) x BC036506 (1-3793)
 Qy 177 AlapheAspSer---AepArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
 |||:::|||||
 ::::

Db 787 GCATGATCTTACCAGCTATTATTTCAGCAACAATCTCAATGTGTTCTCGATGGAGAG 846
Qy 196 SerGlyIleGlySerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 847 AGAGCATCTGGAACAACTATTGTTATAAATCTGGCTGTGTTGAGGTGGATCAAGGGTGAG 906
Qy 216 LeuTrpGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 907 ATGTGGCAGACATGATCTGTACGTCTTACCTCACTTCTCAGCAATAAATCAACAGATG 966
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 967 ACCAACACAGCAGCTGGCTGAGCTAATCGCCAGAGACTGGCTGACGGCCAGCTCCCAT 1026
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspLeu 275
Db 1027 GCAGACATCTGTCTGATCCCAAGAACTCTTTTCATCTCGAGAGACTTGGGCAACATA 1086
Qy 276 GlySerValLeuAsn---AsnAspThrLysLysCysLysAspTrpAlaGluLysGlnPro 294
Db 1087 AGATTTCGAGTTAAATGTCAATGAAGTCTTTGTGTAGTAAACAGCAGCCAGAAAGTTCCC 1146
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 1147 ATTCAGTTCTCCTGGTCAGTTTGTGAAGAGAAAAATGGCTCCAGGCTGCTGCTCCTC 1206
Qy 315 ValThrValArgAspValGlyThrCluLysLeuLysSerGluValValSerProArgTyr 334
Db 1207 ATCTCCTCAAGGCCACACGTGGGAATAATGTAAAGAGTTCTTGAAGAGGTAGATTGC 1266
Qy 335 LeuLeuValArgGlyIleSerGlyCluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 1267 TGCAGGACCTTGCAGCTCTCGAATGGGAAGAG---GAGATATATTAACTCTTTCTTT 1323
Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeu--- 373
Db 1324 AAAGACCCCGCAGAGGGCGCTCGCAGCCCTCCAGCTTGTACATGAGGATGAATACTCGT 1383
Qy 374 AspGlnCysGlnValProAlaValCysSerLeuIleCysValAlaLeuGlnLeuGlnAsp 393
Db 1384 GGTCTGTCGGAGTCGCCATCTTATGTGGATCAGCTGTACTGTCTGAAAGCGGCAGATG 1443
Qy 394 ValValGluGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPhe 413
Db 1444 GACAAGGGCGTGACTTCCAGCTCTGTGTCGCAACACCCACTGATCTACATGCCACTTT 1503
Qy 414 ValPheHisGlnLeuThrProArg---GlyValValArgArgCysLeuAsnLeuGluGlu 432
Db 1504 CTGTGCTGATGCGTTGACATCAGAGCGCTGGACTTACTGCAATCAGTATCACTAGGT--- 1560
Qy 433 ArgValValLeuLysAspPheCysArgMetAlaValGluGlyValTrpAsnArgLysSer 452
Db 1561 -----CTCCTAAACGTCGTGTTGTGCTGCTCAGAGAGACTGTTCTGAGCAGCCCTG 1614
Qy 453 ValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeu 472
Db 1615 AATTTCAGTGTGAAGACCTCAGATGTGTTGGTGTACTGAGGCTGATGTCGTGCTG 1674
Qy 473 PheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHis 492
Db 1675 CAGCGCCGGAATATCTTTTCCGAGCAACACTCATAAAGACCGTTACAAGTTTCATACAC 1734
Qy 493 LeuSerLeuGlnAspPheCysAlaLeuTyrTyrValLeuGluGlyLeuGluIleGlu 512
Db 1735 TTGAACGTCAGGAGTGTGTTGACAGCCATTCGATTTCTG-----ATGGCAGTACCC 1785
Qy 513 ProAlaLeuCysProLeu-----TyrValGluLysThrLysArgSerMetGlu 528
Db 1786 AACTATCTGATCCCTCAGGCAGCAGAGAGTATAAAGAGAGAGAGAACATACTCTGAC 1845
Qy 529 LeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeu 548
Db 1846 TTTAATCAA-----GTGTTTACTTTTCTTTTGTGCTT 1878

Qy 549 ValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGly 568
Db 1879 CTAATGCAACACAGGAGAAAGATTCTTGACATCTCTTTGGATACACAGCTACCGATGGTA 1938
Qy 569 ValLysGlnLysLeuLeuHisTrpValSerLeu-----LeuGlyGlnGln 583
Db 1939 GACAGCTTCAAG-----TGGTACTCGTGGGATACATGAACATTGGACCGTGAC 1989
Qy 584 ProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGln 603
Db 1990 CCGGAAAGTTGAG-----CACCATATGCTTTGTTTACTGTCTCTATGAGATCGG 2043
Qy 604 AspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsn 623
Db 2044 GAAGAAGATTTGTGAAGACGATTGTGGATGCTCTCATGGAGGTACAGTTTACCTTCAA 2103
Qy 624 GlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLys 643
Db 2104 TCAGACAAGATATGATGGTCTCATTAATACTGTCTGGATTACTGTCTACCTGAGGACA 2163
Qy 644 IleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProVal 663
Db 2164 CTAAAGTTGAGTGTTCAGCGCATCTTTCAAAACAAAGAG-----CCACTT 2208
Qy 664 Val---ProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCys 682
Db 2209 ATAAGGCCAACTGCTAGTCAAAATGAAGAGCTTTGTC-----TACTGGAGAGAGATCTGC 2262
Qy 683 SerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThr 702
Db 2263 TCTCTTTTATACAATGGAGACCTCCGGAGCTGCATATCTTTGACAATGACCTTAAT 2322
Qy 703 GluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThr 722
Db 2323 GGTATTTCAGAAAGGATTCTGTCTAAAGCCCTGGAGCAATTCTAGCTGTAACTTCGCACA 2382
Qy 723 LeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMet 742
Db 2383 CTTCAAGTTGCTATGTCTGACTGCTTCTGGTGTGTTGAAGACTTACTCAAGGCTTTGGCT 2442
Qy 743 AlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspVal 762
Db 2443 CGTAATCGGAGCTGACATACCTGAGTATCAACTGACGTGTCATTTCCCTAAATATGTTT 2502
Qy 763 ArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAsp 782
Db 2503 TCACCTTCTGCATGACATCTCTGCACGAGCCACATGCAATAATAGTCACTCTGAGAATACAC 2562
Qy 783 CysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerPro 802
Db 2563 -----AGCCCAACGGCAGCGCTCAG----- 2583
Qy 803 SerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeu 822
Db 2584 -----CCTGGAATCCCAACACTT 2601
Qy 823 SerAspAla-LeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGln 842
Db 2602 TGGGAGCGCGAGCGCGCAGATCACCTGAGCTTGTATGAAA-----TGTGA 2646
Qy 842 YlleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThr 862
Db 2647 TTTGCGAGCCGCGAATGCGAANAATCGCTCTCTCTCATCATGTCGGGGAGTCTGAG 2706
Qy 862 rHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSe 882
Db 2707 AAAAGTACCTTATCCAGCAATCCGCTGAGGAGCGAGCGGATGACATACATACTGTGTGATGC 2766
Qy 882 rMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThr 902
Db 2767 CTTGCTTATCCCAACTGCATCTTATATACCTGCTGTGTAGTCTCTGCTGCTCACTGA 2826


```
QY 902 rlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSe 922
Db 2827 AAATTGCTGCAGCCCTGGAGAGTGCTTCTGTTCAGCCCACTTAAGACCACTAGA 2886
QY 922 rLeuSerMetAsnProValGluAspAsnGlyVallyLeuLeuCysGluValMetArgGl 942
Db 2887 CTGTGTGGAATCGCTTAAANAATACGGAGTGTTCATGTGAGCTTCCCTTCTGCTGT 2946
QY 942 uProSerCysHisLeuGlnAspLeuVallyCysGlyHisLeuThrAlaAlaCysCy 962
Db 2947 TCCAACCTGTCTAGAGGAGCTTCATCTGCTGGCTGCTTCTTTAGCAGCGATATCTG 3006
QY 962 sGluSerLeuSerCysValIleSerArgSerArgHisLeuIlySerLeuAspLeuThrAs 982
Db 3007 TCAATATATATGCTAGTATTTGCTACTAATAAAGAACTGAGAGCTGGAGATTGGGAG 3066
QY 982 pAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuIlyAsnSe 1002
Db 3067 CACAAAATAGAGATGACGAGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3126
QY 1002 rValLeuThrArgLeuGlyLeuAlaCysGlyLeuThrSerAspCysGluAlaLe 1022
Db 3127 CATGTTGGTATATTTGGCTAGAGAGTGCATGTTAACAGTGCCTGCTGTGATCTCT 3186
QY 1022 uSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnValGlnAsnAspH 1042
Db 3187 TGCCTCTGCTTCTTACCACCAACAACTGAGAAAGACTCAACTGCTTCAAAATCACTT 3246
QY 1042 eSerProIlyGlyMetMetLeuLeuCysSerAlaPheAlaCysProThrSerAsnLeuGl 1062
Db 3247 GGGCAATGATGAGTGTGCAAACTCTTGGAGAGCTTGATGACCCAGATGTGTACTTAA 3306
QY 1062 nIleLeuGlyLeuThrPlyTrpGlnTyProValGlnIleArgLysLeuGluGluVa 1082
Db 3307 GGTAGTGGCTTCATTAAGTGGCTGACACACACAAACCCAGCAGTGTGTGACACTGT 3366
QY 1082 lGlnLeuLeuIlyProArgValValIleAspGlySerTrpHisSerPheAspGluAspAs 1102
Db 3367 AAAGGAAGAAACCCAGTTGATCTTCTGTCGAACTTGTGCTTTTAAAGGAAGGC-- 3424
QY 1102 pArgHisIlyLeuGlyLeuThr 1109
Db 3425 ----AGAGAAATTTGGTGACA 3442

RESULT 9
AK054426
LOCUS
DEFINITION
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330024M09 product:hypothetical
RNI-like structure containing protein, full insert sequence.
ACCESSION
AK054426
VERSION
AK054426.1 GI:26344236
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
```

Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 3275)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. 3275
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E330024M09"
/db_xref="WGI:2426593"
/db_xref="taxon:10090"
/clone="E330024M09"
/sex="female"
/tissue type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
169. .2760
/note="unnamed protein product; hypothetical RNI-like structure containing protein (SCOP|52047, evidence: SCOP) putative"
/codon_start=1
/protein_id="BAC35775.1"
/db_xref="GI:26344237"
/translation="MASLFSDFGFIVYWKELNKFMYFKELLHLEILQMLQKISWT EVKEASREDLAILLVKHCQGNQAWDTTFRVQFMIGRNVITNRTATGEIAHSTIYRAHL

FEATURES

source

CDS

Qy 571 GlnLysLeuLeuHisTrpValSerLeuLeu-----GlyGlnGlnProAsnAlaThrThr 588
Db 1732 CGACAGTGTGTTTGTGGCTGGAATCTCTATTGGACACTTTACATCTCGACCTAAAAAA 1791
Qy 589 ProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheVal 608
Db 1792 ATA---AATACCATGAAGTCTTTTACTGCTGTTTGAGATGGAAGAGAGTCTTTGTA 1848
Qy 609 ArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeu 628
Db 1849 CAGTCAGCAATGAAGTGTAGGGAACAGATGACGTGTGGTTAGGGTTATTCTGATTTT 1908
Qy 629 IleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspVal 648
Db 1909 ATTGTGTGCTGCTACTGCTTAAGCCTGCTCTGACCTGACAGACTCTTCATCTTCAGCT 1968
Qy 649 LysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrpMet 668
Db 1969 CAAATGTGCTGAATGAAGAG-----CTAGGCCAG 1998
Qy 669 ArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHis 688
Db 1999 AGGGGAAGACTACTCAT-CTTTGGCATCAATATGCTCTGTGTTCTTAAGGAAT 2052
Qy 689 ProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThr 708
Db 2053 AAAGATATCAAAACACTCAGATGAAGACACTATTTTCAATGACACCACTGTTTAAATC 2112
Qy 709 LeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAla 728
Db 2113 TTTTATCTTATTGAAGACTCCAGCTGATCTCTTAAACACTTGTGGCATATAATGTT 2172
Qy 729 GlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArg 748
Db 2173 TCTTTTATGTGAT--AAGCGCTATTCTTGGAAATGATTCAAGCTACATTTGGAG 2229
Qy 749 SerLeuAsnLeuGlyGlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAla 768
Db 2230 GAATGTATCTCAGAGGCACATCTCTCCACAGTGTATGGAATGTTGTGTATATC 2289
Qy 769 LeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHis 788
Db 2290 TTGAACGAGGAGATGCAATAGCGATCTAGTATTAGTAATCTCTCTCTCTGTGAG 2349
Qy 789 AlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSer 808
Db 2350 CATCTCTGGGACTACTTCTGATCTTCTAGGAGACAATCTCTGAGTACTTAAC 2409
Qy 809 LeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVal 828
Db 2410 ATCAGCTACAATAACCTGAAGGTGAAGGCTCTCTGTAGAGCTCTGACTCTC 2469
Qy 829 SerGlnCysAlaLeuGlnLysIleLeuGluAspCysGlyIleThrAlaThrGlyCys 848
Db 2470 CCAAACTCTGCTGCTGCTACTAGTTTGAAGCCTGCTCAACTGACTGTGCTGTGT 2529
Qy 849 GlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsn 868
Db 2530 AAGGACCTTGCTTCTTACTTACCGGATATAATGCTGCGGAGGATCAACTGCGGAG 2589
Qy 869 AsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCys 888
Db 2590 AACTCTCTGGGCTTCAGTGGATTTGTTGTGTGTGTAAGGCTATCAAGACCAACGCTG 2649
Qy 889 SerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeu 908
Db 2650 ACCTTGTATGAGTAAACTCGAATGCTGACTTTGTATAGCGATTCCAGGAATTTCTA 2709
Qy 909 Ala-Leu----- 910
Db 2710 CTTCTGAGATGGAAGAAATAAATCTTGAGCATCGAAATGTTGTGTAGGAGAGCA 2769
Qy 911 -----AlaLeuMetGlyAsnSerTrpLeu---ThrHisLeuSerLeuSe 924

Db 2770 GAAATCAAGGTGGACATCTCGATGCAAGAACATGATGATTTGTGAACCTCAAACTACT 2829
Qy 924 rMetAsnPro---ValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluPr 943
Db 2830 TCTGNAATGAAGAGAGAGGATCTTTATTGTTGCTTATTGTTGTAATAATACAGTCACT 2889
Qy 943 oSerCysHisLeuGlnAspLeuGluLeu 952
Db 2890 AACACTGCAGACAGCAGGATTCTTTTGTG 2917
RESULT 10
AK054264 3400 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
DEFINITION full-length enriched library, clone:E330007A02 product:Weakly
similar to PAN1, full insert sequence.
AK054264 GI:26344106
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
2
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PubMed 11042159
3
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Fujimori,S., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuda,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PubMed 11076861
4
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,K., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome-gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers

1. 3400
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:E330007A02"
 /db_xref="MGI:2426301"
 /db_xref="taxon:10090"
 /clone="E330007A02"
 /sex="female"
 /tissue_type="ovary"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days pregnant adult"
 76. 2553
 /note="unnamed protein product; Weakly similar to PAN1 (SCOP|52047, evidence: SCOP)
 putative"

/codon_start=1
 /protein_id="BAC35710.1"
 /db_xref="GI:26344107"
 /translation="MEHFDPLGNLKDVLRLKDLTELNNKRTLRSLPDTWKQINK
 LTMDLANGAQLAELTHDCPSAWIKRTVQVLEINRVDLAELVQKIEAVLKVPBE
 KVSKEPSPGTLFFPVQAGKRPDKQKEWKTRYTAQKQNFMPKCNKSIYVVT
 ESKYTLANKPKTETFPFAHAIVLHPPGSGKTTMAQLMWSKQAIIFSCAFYI
 SCRVNNTKPCFTFAHLSMDNPSWRDVIKLEFLFVVDGDFDLTPAGALIRD
 LCGDNTVQKPEVLLGSLKKMAPHATLLVTRTOSLHOFVMDQPLVETLGLFLE
 QKEQYKQYFEDSEGEEDKGEKALRKALKEVCNADLYQMASLPTACGIFCLLEL
 RMKGDELSTLCQYITSMFLNFCVFSSTCEDHNEEFQILFKKICILAAASLLSQ
 VPIICEEDFLTKQNLNHLNHPVCRHILFKDSSSTHCLFICIGIQQLAAIIFVQEL
 QGESKVSQKYSIQNLSREARLKNPDLGSLLPVFGILNETRIQELKTFPGCQISTEV
 KRKLECEGSENKPLLLMMQELISCLYSEBGFVKEAMVLPEDISLHKTSTDI
 HASCLKNQNLQTMSLKVEKAVPENVAALESFAKHQSPDRMLTFTWDFCDTFN
 SNKLVLIDTHESFLNSALEILCEKLPASACCLQKVLKNI SPDDAYEKLCLIFNGY
 KTI SHLLIQGNLDSMHSEVULKNPACNLKFLUSLGSCTSAQAQRWDDFFFLVKVQNS
 LIFLDLSDNLSLDSKALLCNWKEPK"
 3383. 3388
 /note="putative"
 3400
 /note="putative"

ORIGIN
 Alignment Scores:
 Pred. No.: 9,47e-84 Length: 3400
 Score: 940.00 Matches: 275
 Percent Similarity: 46.58% Conservativity: 208
 Best Local Similarity: 26.52% Mismatches: 477
 Query Match: 13.29% Indels: 78
 DB: 11 Gaps: 19

US-10-066-521-6 (1-1344) x AK054264 (1-3400)
 Qy 12 TyrGlyLeuGlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPheLysGlu 31
 Db 100 TTCAACCTAAAGACGCTCTGAGAAAGCTGACAAAGACTGAGTTGAACAATTTCAAGAGA 159

Qy 32 LeuLeuLysLysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGlu 51
 Db 160 ACATTGAGGAGCTGTCTCTGCCAGATACATCAACAGCATCAACAGCTGACATGCAG 219
 Qy 52 AsnAlaAsnValGluCysLeuAlaLeuLeuLeuHieGluTyrTyrGlyAlaSerLeuAla 71
 Db 220 CTGGCTAATGGAGCTCAGCTAGCAAAATCTCACAGACCCTGCCAGCTGCTGGATA 279
 Qy 72 TrpAlaThrSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAla 91
 Db 280 AAGAGATCACCTCCAGATCTTGAAGAAATAAACCGTGTAGATCTAGCTGTAGTA 339
 Qy 92 ArgAspMetLysLysIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGlu 111
 Db 340 -----GTAAGCAATCGAAGAGCGATTAAGGTACCTGAGGAAAAGTTTCT 390
 Qy 112 ThrGluGluGlnGluIleSerGlnAlaMetGlu-----GlnGluGlyAla 126
 Db 391 TCTAAACCAAGAGACCTCGGGGACACTAATCTCCCATGGAATTTTGTTCAGGAGCT 450
 Qy 127 ThrAlaAlaGluThrGluGluGlnGlyHieGlyAspThrTrpAspTyrLysSerHis 146
 Db 451 AAAAGCCAGAGAGCAAGCAAAAAGAG-----GAATGGAAGACTAGA 492
 Qy 147 ValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAsnThrAlaAla 166
 Db 493 TACACAGCTAATGG-----AAGCAGAACTTC----- 519
 Qy 167 AspTyrPro-----GluMetGlnThrLeuAlaGlyAlaPheAspSerAspArg 182
 Db 520 ---TGGCCTAAATGCAATAAAGAAATCTATGTGTTACCGAGAGCTACAAGACACTCCTG 576
 Qy 183 TrpGlyPheArgProArg-----ThrValValLeuHieGlyLys 195
 Db 577 GCATGTGTAAACCAAGATAGAAACCCATTTGCTCATGCAATTTGTGTCATGCCCT 636
 Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
 Db 637 CCAGGCTCTGGGAAACCAACCAATGGCAAGCAGCTAATGTTAGAAAGTGCAGAGATAAG 696
 Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
 Db 697 CAGGCCAG---ATTTCCTCATGTGCTTCTACATCAGCTGCAGGGAAGTAAATAACT 753
 Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
 Db 754 AAACCTTGCACTTTGCCACCTCTCCATGAGCAATCCCTCTGAGGAGGACTGTGTG 813
 Qy 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeu 275
 Db 814 ATTCGAGATCTAATCTCTGGGAAAGGAATCTTATTTGTTGTTAGTATGTTGTTAGCTG 873
 Qy 276 -----GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
 Db 874 ACATTCCAGCAGGAGCTCTGATCCGTGAC-----CTTTGTGTGACTGCAACACA 924
 Qy 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
 Db 925 GTGAAGCCTGTAGAGCTCTTACTGGCAGTTTGTCTGAAGAGGAAGTGCACCTCATGCC 984
 Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
 Db 985 ACCCTGCTGTATACCACGACGACAGTCTTTGCACCAATCTTTCTCATGATGGATCAG 1044
 Qy 332 ProArgTyrLeuValArgGlyIleSerGlyGluGlnArgIleHieLeuLeuLeu--- 350
 Db 1045 CCACTACTAGTAAACTCTGGGCTTCTTGGAGCAAGAAAGAGCAGGAGTACTTTCAGAAA 1104
 Qy 351 -----GluArgGlyIleGlyGluHieGlnLysThrGlnGly 362
 Db 1105 TACTTTGAAGATGAGAGGGGTGAGGAGGAGTAAAGGTGAGGGAAGGACCTTAGAGCT 1164
 Qy 363 LeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGly 382

```
Db 1165 CTCAAGGAAGTGAGATGCAATGCTGATCTATACCAAAATGGGCTCACTTCTCTACTGCAATG 1224
Qy 383 SerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerValAlaProPhe 402
Db 1225 GGGATATTCTGCTCTGCTGAGCTAGGATGAAGAGGGGGAAGACCTGCTCTGACA 1284
Qy 403 AsnGlnThrLeuThrGlyLeuHisAlaPhe-----ValPheHisGlnLeuThr 419
Db 1285 TGTCAAGATACACCTCTATGCTTCTGAACTTCTCTGCGAAGTCTTCTCATCAAAACC 1344
Qy 420 ProArgGlyValValArgCys-----LeuAsnLeuGluGluArgValValLeu 436
Db 1345 -----TGTGAGGACCATCTCAATGAGGAATCCAAATCTTATTC 1383
Qy 437 LysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGly 456
Db 1384 AAGAAATATGATCTCTGGCTGCCAATAGTCTCTTGGAAAGGTCCTCATCTCTGTGAA 1443
Qy 457 AspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsn 476
Db 1444 GAAGACTTCTTGACATTAATAATTGAACCTAAACAATTTGCATCCCATGGTGTGCAGGCAC 1503
Qy 477 IleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGln 496
Db 1504 ATCCCTCTTTAAGGACAGTTCCAGTACACACTGTCTCTTTTCATCTGCTCGGTATCCAG 1563
Qy 497 AspPheCysAlaAlaLeuTyrTyrValLeuGlu---GlyLeuGluIleGluProAlaLeu 515
Db 1564 CAACTCCTGGCTGCAATATATTGTTTCAGGAGCTTGACAGGAAAG-CAAGGGTGTCTC 1622
Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db 1623 TAAATATAGTATACAGAACATGCTTTCAAGGGAAGCAAGATTAAAAACCC----- 1673
Qy 536 IleHisSerLeuTrp-MetLysArgPheLeuPheGlyLeuValSerGluAspValArgAr 555
Db 1674 ---TGACCTGTCTGGGCTGTGCCATTGATTTGGCCCTCGTGAATGAGACTCGTATTCA 1730
Qy 555 gProLeuValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHi 575
Db 1731 AGATTGAAGACTACTTTGGCTGCCAATATCAACTGAGGTCAAGAGAAATTCCTGGA 1790
Qy 575 eTrpValSerLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAl 595
Db 1791 GTGTGAGTCA-----GGGAAATAAACCCCTTACTATTACTGATGATATGCAAGAAAT 1844
Qy 595 aPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPh 615
Db 1845 CTGTGCTCTCTATGAATCTCAGAGAGGGGTTTGTGAAGGAAGCCATGGTCTCTT 1904
Qy 615 eGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLe 635
Db 1905 TGAGGACATCTCTTGGCATTTGAAACACGAGCACTGACCTCATTCATGCTTATTCGCT 1964
Qy 635 uGlnHisCysProTyrLeuArgLysIleArgValAspVal---LysGlyIlePheProAr 654
Db 1965 CAAGAACTCTCAGAACTTGCAGACAATGTCCTGGAAGTGAAGAGGGCTTTTCCAGA 2024
Qy 654 gAspGluSerAlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIle 674
Db 2025 GAATGTGTGCTGCACTAGAGTCCACAGTCAAGCATCAAAGGTCCCGGATGAACAACGCAT 2084
Qy 674 eGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLe 694
Db 2085 GTTAATCTTCTGGACAGACTTTTGTGACACATTTAATTCAAATAAGAAAGCTGGTATTCT 2144
Qy 694 uAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuAr 714
Db 2145 AGACATCCATGAAGAGCTTCCTCAACAGCTCTGCATCTGGAGATTCTTTGTGAAGAGCTCC 2204
Qy 714 gHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyVa 734
```

RESULT 11

```
BM454123 1119 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6419582 NCI CGAP Ov44 Mus musculus cDNA clone
DEFINITION IMAGE:5504420 5', mRNA sequence.
ACCESSION BM454123
VERSION BM454123.1 GI:18503163
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1119)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
```


10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6

(bases 1 to 2730)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

7

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

8

Location/Qualifiers

1. 2730

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:A330055K17"

/db_xref="MGI:2404003"

/db_xref="taxon:10090"

/clone="A330055K17"

/sex="male"

/tissue type="spinal cord"

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6

(bases 1 to 2730)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

7

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

8

Location/Qualifiers

1. 2730

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:A330055K17"

/db_xref="MGI:2404003"

/db_xref="taxon:10090"

/clone="A330055K17"

/sex="male"

/tissue type="spinal cord"


```
/clone.lib="RIKEN full-length enriched mouse cDNA library"  
/dev stage="adult"  
1..2730  
/note="similar to VASOPRESSIN RECEPTOR [Rattus norvegicus]  
(SPTR|063035, evidence: FASTY, 82.2%ID, 100%length,  
match=1440)"
```

misc_feature

ORIGIN

Alignment Scores:

```
Pred. No.:      3,46e-56      Length:      2730  
Score:          669.50      Matches:      235  
Percent Similarity: 42.53%      Conservative: 132  
Best Local Similarity: 27.23%      Mismatches:  349  
Query Match:      9.46%      Indels:      149  
DB:              11          Gaps:       21
```

US-10-066-521-6 (1-1344) x AK039531 (1-2730)

```
QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArg 205  
DB 14 CAGCGCTGACGGTGGTCTCGAGGCCCGCGCATCGCAAAACCATGGCGCGCAAG 73  
QY 206 ArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225  
DB 74 AAAATCCTCTATGACTGGGCAGCAGGTAAGCTATACACACGCCAGGTGGACTTTGCCCTTC 133  
QY 226 PheLeuProValArgGlu---MetGlnArgLysLysGluSerValThrGluPheIle 244  
DB 134 TTCATGCCGTGGCGAGCTCTCGAGCGCGCGGAAAGCGCAGCTTGGCGGACCTGGTG 193  
QY 245 SerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArg 264  
DB 194 CTGACACAGTCCCGGACCGCGCTTGGCCAGTGAAGCGGATCTATGGCGCAACCCAAATCGC 253  
QY 265 LeuLeuPheIleIleAspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLys 284  
DB 254 CTACTGTTTCCTCGGAGCGTGCAGACGAGCTGCTACTTTACCATCTCTCAGAGCCCACT 313  
QY 285 LeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArg 304  
DB 314 CCCTGCAAGACCCCTTTGGAGGCTACAAAGTGGCTTGAGAGTGTCTGAGCGGGCTGCTGAGC 373  
QY 305 LysValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLys 324  
DB 374 CAGAGCTGTTTACAGAGCTCGCTTCTAGTGACTACAGCCCATGCCCTACAGGGAGG 433  
QY 325 LeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGln 344  
DB 434 CTGACGGGCAGATTGTGCTGCCACAGTGCAGCAAAATACCGCGCTTCTCAGACAAAGAC 493  
QY 345 ArgIleHisLeuLeuLeuArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArg 364  
DB 494 AAAAAAAGATATTTCTTCAAGTTCTTCGGGACGAGAGGAAGGACGAGCGCGCTTACCCGC 553  
QY 365 AlaIleMetAsnAsnArgGluLeuAspGlnCysGlnValProAlaValGlySerLeu 384  
DB 554 TTCGTGAAGAACAGACAGACCTCTTTCAGTGTGCTTGTGCCCCCTCTGCTGCTGGATC 613  
QY 385 IleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGln 404  
DB 614 GTGTGCACAGTGTGCAGCAGCAGCTAGAGCTGGCGGGGTCTCTCTCGTACCTTCCAAG 673  
QY 405 ThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValVal 424  
DB 674 ACCACTACTTCTGTGTACTCTCTTCAATCACCAGCATGTGTGAAGTCTGACGAGCCACCAAT 733  
QY 425 ArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaVal 444  
DB 734 GGACCCCGGTTTCAGGGAGAG-----CTGAGAACGCTGTGTCTGCTGGCCAGG 781  
QY 445 GluGlyValTrp---AsnArgLysSerValPheAspGlyAspLeuMetValGlnGly 463  
DB 782 GAGGGCATCTGGATCATATAAAGCAAAATTTCTCAGAAAGAGGACCTGGAGAAATTCGAAG 841
```

```
QY 464 LeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAsp---Ser 482  
DB 842 CTTTCGGGGTTCCCAAGTTTCAGACAATATTTCTCAACAAGAAAGAGATATACCGAGGTGCTA 901  
QY 483 HisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeu 502  
DB 902 AAAACTGAGGTCACTACCGTTTCATTGACACAGAGCTTCCAGGAGTTCTTGGCTGCATTG 961  
QY 503 TyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLys 522  
DB 962 TCATACCTACTAGAAAGT---GAGCGAACTCCCGGAGCGCCCGCAGCGGTGTGCAGAAG 1018  
QY 523 ThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLys 542  
DB 1019 CTCCTGAACCTCT-----GACGGGAGCTAGCTGTGTCATCTTGCACTGACCCACC 1066  
QY 543 ArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGly 562  
DB 1067 CGATTCTCTTTGGCTCTGTAATACAGAGGGCTTCGTGACATTGGAAACCAATTTTGGC 1126  
QY 563 CysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGln 582  
DB 1127 TGTGTGGTGCACATCATGTGAAAGAGCACCTCTCGGTGGGTACAA-----GGACAG 1180  
QY 583 ---GlnProAsnAlaThrThrProGlyAspThr----- 592  
DB 1181 AGCCACCCCAAGGGGCCACCAAGTAGGGGCAAAAAGAGACTGCTGAGCTGGAGGACATCGAG 1240  
QY 593 -----LeuAspAlaPheHis 597  
DB 1241 GAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1300  
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617  
DB 1301 TGCCTGTGTGAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1360  
QY 618 ValTrpLeuProIleAsnGln-----AsnLeuAspLeuIleAlaSerSerPheCysLeu 635  
DB 1361 ATAGTACTGGAGCGAGTCTGAGTTGACCCGCTGAGACCTTGAAGTTCTGAATTTACTCGGTG 1420  
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655  
DB 1421 CAGTGTGCCAGATGGCAGGCTCTGAGACTGTGGCTGTGGCTGTGGCGGCAAG 1480  
QY 656 Glu----- 656  
DB 1481 GAGAAAGAAAGAAAGAAAGAGCCTGTGAGCGCTGAAAGGGCTCTCAAGACCAACCAAG 1540  
QY 657 -----SerAlaGluAlaCys 661  
DB 1541 AAACAAACCCCAAGTCTCTTGTGCTGCCACTCTGTGAGACATGACTACCCCGAAATGT 1600  
QY 662 ProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAspPhe 681  
DB 1601 CATCTGAGTGTCTGATCTTGTTCACACTGCAGACTCCCTGATGCAGTTTCCCGAGACCTT 1660  
QY 682 CysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeu 701  
DB 1661 TCCGAGGCGCTGAAGGTAGTCTCTGCGCTTAAGGGAGTTGGCGCTCTCCAGAGCCCGCTC 1720  
QY 702 ThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGln 721  
DB 1721 ACCAACACAGCGCTGCGTTTACTGTGCGAAGGCGCTGGCTGGCCCAAGTCCAGGTTGAAG 1780  
QY 722 ThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleVal 741  
DB 1781 ACCTCAGGATGAGCTGCTGACCTCCAGAGAGGTGATCAATTACTTGTGTCATCGTGCTC 1840  
QY 742 MetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAsp 761  
DB 1841 CAGCAGAGCCAGTCTCCACCACTCTGGACCTCAGTGGCTGTGAGTGTGCTGCTGCTGCTCAT 1900
```



```

Qy 762 ValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeu 781
Db 1901 GTGGAACCTCTGTGTGCGACCCCTGAAGCACCTAAATGCTCCCTAAAGACCTCAGTCTG 1960
Qy 782 AspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrSer 801
Db 1961 ACCTCTGGAGCTGAGTGAGACTACTAGGGATCTTCAAGCTGTGAAGACATCAAG 2020
Qy 802 ProSerLeuLys-SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetPr 821
Db 2021 CCAGATCTGTCCATCATATTTCAAAATAAGGCACACCTCAGCGCCCTCGA----- 2072
Qy 821 oLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCy 841
Db 2073 -----AGGGTGAGGACGTGT----- 2087
Qy 841 sGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLe 861
Db 2088 -----CTTCTGAAGCTCCGGAGTCC 2107
Qy 861 uThrHis-----LeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLe 879
Db 2108 AGAACACGATGAAGATGCTGTAGCCCAAGTCCCTCTG-AAT-----ATGACCTCCCC 2160
Qy 879 uCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMet----- 894
Db 2161 ATTCAAAAGCAAGAGATGCCAGC-TTCTCCCTCCAGGCAAGTCTTTCAAGCCACAGAAG 2219
Qy 895 -----LeuAsnGlnCysHisLeuAspThrAlaG 904
Db 2220 GCGAAGCAAAAGACCCAGGTAGATACAGCGCCTTGATAGCTCTCGGTAGATAGCTGGCAG 2279
Qy 904 y-----CysGlyPheLeuAlaLeuAla---LeuMetGlyAsnSerTr 917
Db 2280 TGCCCTGGCCTGGACACTCTGGGTGTGCTTCCCTCAGTCCACACACCAATGCCTT 2339
Qy 917 pleu-----ThrHisLeuSerLeuSerMetAsnProValGluAspAsnGly----- 932
Db 2340 CTTTCTGAGCCTACCATCTCTGCTCCCTGACGATATCCCAACAGTAATAGAAGCATTGAG 2399
Qy 933 -----ValLysLeuLeuCysGluValMetArgGluProSerCy 945
Db 2400 CAGCTGACGGTGTGCTGAATGTTGTCACTATGTTGTATCTG-----GCCCCAGCTG 2453
Qy 945 sHisLeu 947
Db 2454 TCATATG 2460

RESULT 14
LOCUS AY407369
DEFINITION Mus musculus RNH gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY407369
VERSION AY407369.1 GI:39763340
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1371)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003).
14671302
2 (bases 1 to 1371)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

```

```

Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
source
1..1371
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1371
/gene="RNH"
/locus_tag="HCM2867"

ORIGIN
Alignment Scores: 4.27e-54 Length: 1371
Pred. NO.: 643.50 Matches: 150
Score: 52.31% Conservative: 76
Percent Similarity: 34.72% Mismatches: 197
Best Local Similarity: 9.10% Indels: 9
Query Match: 29 Gaps: 3
DB:

US-10-066-521-6 (1-1344) x AY407369 (1-1371)
Qy 659 GluAlaCysProValValProLeuTrp-----MetArgAspLysThrLeuIle 674
Db 49 GAGCTCTTCTCTGATCCAGCAATACGAGTGGTCAAGTGGATGCTGTGGCTCACT 108
Qy 675 GluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeu 694
Db 109 GAAGTGGGTGCAAGACATCAGCTCAGCAGTCCAGCAACCTGCCCTGACAGAGCTC 168
Qy 695 AspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArg 714
Db 169 AGCCTTCGCAACCAAGAACTGGGTGATGGGTGCTGTGGTCTCCAGGGCTCGAG 228
Qy 715 HisProThrCysValIleGlnThrLeuMetPheArgAsnAlaGlnIleThr---ProGly 733
Db 229 ATCCCACTTGTAAATGCCAGAGCTGAGCCTTCAGAACTCGCGCTTGACGAGGCTGGC 288
Qy 734 ValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGly 753
Db 289 TGTGGGATCTGCTCGCATGCTGGCTCTTTGTCTACCTGCGTGAGCTACATCTCAAT 348
Qy 754 GlyThrHisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLys 773
Db 349 GATAACCTATGGGGGACGCGAGCCTGAAGCTGCTCTGTGAAGGACTTCAGGACCCGAG 408
Qy 774 CysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLys 793
Db 409 TGGCGTCTTGAGAAGCTTCAGTTGGATATCTGTAACCTCAGCTACAGCTACGAGTGGAGCCCC 468
Qy 794 IleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLys 813
Db 469 CTGCTTCAGTGTCTCAGGGTGAAGCTGACTTTAAAGGAGCTAGTATTGAGCAACACGAC 528
Qy 814 ValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeu 833
Db 529 CTCCTCAACCTGGTGTCCGTATTCTGTGCGAGGCTGGAAGGACTCTGCTGTCAACTG 588
Qy 834 GlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSer 853
Db 589 GAGTCACTCAAGCTGGAGAACTGTGGTATCACAGCAGCCCAACTGCAAGGATCTGTGGCAT 648
Qy 854 AlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeuGlyAsn 873
Db 649 GTTGTGGCTCCCAAGCTTCGCTACAGGAAGTGGACTTGAGCAGCAACAGCTAGGTAAT 708
Qy 874 GluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeu 893
Db 709 GCAGGCATTCAGCACTGTGCCAGGAGTCTGTCTCCAGCTGTAAGCTCAGGACTCTG 768

```



```
QY 942 GluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCys 961
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 GAACCTACTGTGAC-CTTCAGAACTGGAACTAGTGGACTGCCAACTCACACAGAACTGC 124
QY 962 CysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThr 981
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 TCGGAGGACCTGGCTGTATGATCATACAAACCAAGCACTTAAAGTTTGGATCTTGGT 184
QY 982 AspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsn 1001
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 AACAAAGCCCTGGGTGACAAAGAGTCTATAACCCCTGTGTGAGGACTGAGCAAGTAGC 244
QY 1002 SerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAla 1021
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 AGCTCCCTGAGGAGACTTGGGTGGGGGCATGTAAGTTGACTTCCAATTGCTGTGAGGCA 304
QY 1022 LeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsn 1041
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 TTGTCTATTGGCCATCTCTTGCACCCCTCACCTGAACAGCCCTAAACCTGGTGAAGAATGAC 364
QY 1042 PheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeu 1061
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 TTCAGTACATCGGGATGTTGAAGCTGTGCTCTGCTCCAAATGCCCTGTCTCTAACCTG 424
QY 1062 GlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuGluGlu 1081
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 GGGATAATTGGCCTGTGGAAGCAGGAGTACTATGCCCGAGTGAGAGACAGCTGGAGGAA 484
QY 1082 ValGlnLeuLysProArgValIleAspGlySerTrpHisSerPheAspGluAsp 1101
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 GTTGAGTTTGCAAGCCCAACGCTGGTGATTGATGGTGATTGGTATGCTAGTGATGAAGAT 544
QY 1102 AspArgHis 1104
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
545 GACCGAAAC 553
```

Search completed: July 16, 2004, 06:23:48
Job time : 7508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 09:47:35 ; Search time 198 Seconds
(without alignments)
3766.939 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGVLVS 1344

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -DB=us-10-066-521-6
-O=/cgn2_1/USPTO.spool/US10066521/runat_13072004_122318_11560/app_query.fasta_1.1543
-DB=Issued_Patents_NA -QFMT=fascap -SUFFIX=Oligo.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn 1.1.107 @runat_13072004_122318_11560 -NCPUs=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:
5: /cgn2_6/ptodata/2/ina/PTC COMB.seq:
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	0.6	29	4	US-09-100-703A-19
2	8	0.6	423	4	US-09-621-976-16796
3	8	0.6	426	4	US-09-621-976-16797
C 4	8	0.6	549	4	US-09-252-991A-5717
5	8	0.6	618	4	US-09-621-976-788
6	8	0.6	805	3	US-08-961-083-207
C 7	8	0.6	805	4	US-09-536-784-207
8	8	0.6	855	4	US-09-489-039A-2718
9	8	0.6	891	4	US-09-252-991A-5722
C 10	8	0.6	944	4	US-09-262-856A-13
C 11	8	0.6	958	2	US-08-632-514C-9
C 12	8	0.6	958	3	US-09-188-177-9

C 13	8	0.6	1117	4	US-08-858-207A-40	Sequence 40, Appl
C 14	8	0.6	1126	4	US-08-961-527-13	Sequence 13, Appl
C 15	8	0.6	1176	4	US-09-491-577-75	Sequence 75, Appl
C 16	8	0.6	1272	4	US-03-489-039A-6621	Sequence 6621, Ap
17	8	0.6	1278	4	US-03-252-991A-5750	Sequence 5750, Ap
18	8	0.6	1299	4	US-09-489-039A-6764	Sequence 6764, Ap
19	8	0.6	1566	4	US-09-833-381-1975	Sequence 1975, Ap
20	8	0.6	1599	4	US-09-252-991A-13921	Sequence 13921, A
C 21	8	0.6	1622	3	US-09-180-939-2	Sequence 2, Appli
C 22	8	0.6	1750	4	US-09-262-856A-7	Sequence 7, Appli
C 23	8	0.6	2155	4	US-03-371-338-8	Sequence 8, Appli
C 24	8	0.6	2280	4	US-09-252-991A-13535	Sequence 13535, A
25	8	0.6	2866	4	US-09-724-864-30	Sequence 30, Appl
C 26	8	0.6	3318	4	US-09-540-236-1379	Sequence 1379, Ap
27	8	0.6	3336	4	US-09-717-364A-26	Sequence 26, Appli
C 28	8	0.6	3393	3	US-09-104-324B-1	Sequence 1, Appli
C 29	8	0.6	3393	3	US-09-162-713-1	Sequence 1, Appli
C 30	8	0.6	3446	4	US-09-332-714-13	Sequence 13, Appl
C 31	8	0.6	3665	4	US-09-976-594-722	Sequence 722, App
32	8	0.6	4086	4	US-09-702-705-1801	Sequence 1801, Ap
33	8	0.6	4086	4	US-09-736-457-1801	Sequence 1801, Ap
34	8	0.6	4086	4	US-09-671-325-1801	Sequence 1801, Ap
C 35	8	0.6	5228	4	US-09-428-711A-15	Sequence 15, Appl
C 36	8	0.6	5418	3	US-09-180-939-1	Sequence 1, Appli
37	8	0.6	5578	1	US-08-081-610-2	Sequence 2, Appli
C 38	8	0.6	5759	4	US-09-898-361-3	Sequence 3, Appli
39	8	0.6	8561	3	US-09-112-450-3	Sequence 3, Appli
40	8	0.6	8561	4	US-09-419-291A-3	Sequence 3, Appli
41	8	0.6	8876	4	US-08-961-527-170	Sequence 170, App
42	8	0.6	26016	4	US-09-326-480A-1	Sequence 1, Appli
43	8	0.6	26270	4	US-09-717-364A-1	Sequence 1, Appli
44	8	0.6	40085	4	US-08-311-731A-26	Sequence 26, Appl
45	8	0.6	50000	4	US-09-146-053-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-100-703A-19
; Sequence 19, Application US/09100703A
; Patent No. 6610300
; GENERAL INFORMATION:
; APPLICANT: SEGERS, Ruud PAM
; APPLICANT: WATERFIELD, Nicolas R
; APPLICANT: FRANDSEN, Peer L
; APPLICANT: WELLS, Jeremy M.
; TITLE OF INVENTION: COLISTRIDIMUM PERFRINGENS VACCINE
; FILE REFERENCE: 97288 US
; CURRENT APPLICATION NUMBER: US/09/100,703A
; CURRENT FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: EP97201888.1
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Clostridium perfringens
US-09-100-703A-19

Alignment Scores:
Pred. No.: 8.25 Length: 29
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0
US-10-066-521-6 (1-1344) x US-09-100-703A-19 (1-29)
Oy 1023 SerleuAlaLeuSerCysAenArg 1030
Db 6 TCTCTAGCTTTGAGCTGTAATAGA 29

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5717
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5717

Alignment Scores:
Pred. No.: 157      Length: 549
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-252-991A-5717 (1-549)
QY 849 GlnSerLeuAlaSerAlaLeuVal 856
Db 109 CAGAGCTGGCATCGGGCTGGTG 86

RESULT 5
US-09-621-976-788
; Sequence 788, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 306..617
US-09-621-976-788

Alignment Scores:
Pred. No.: 176      Length: 618
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-621-976-788 (1-618)
QY 904 GlyCysGlyPheLeuAlaLeuAla 911
Db 26 GGCTGTGGGTTTGGCTCTTGCA 49

RESULT 6
US-08-961-083-207/c
; Sequence 207, Application US/08961083
; Patent No. 6159469

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16796
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16796

Alignment Scores:
Pred. No.: 121      Length: 423
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-621-976-16796 (1-423)
QY 904 GlyCysGlyPheLeuAlaLeuAla 911
Db 48 GGCTGTGGGTTTGGCTCTTGCA 71

RESULT 3
US-09-621-976-16797
; Sequence 16797, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16797
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16797

Alignment Scores:
Pred. No.: 122      Length: 426
Score: 8.00        Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.60%      Indels: 0
DB: 4              Gaps: 0

US-10-066-521-6 (1-1344) x US-09-621-976-16797 (1-426)
QY 904 GlyCysGlyPheLeuAlaLeuAla 911
Db 55 GGCTGTGGGTTTGGCTCTTGCA 78

RESULT 4
US-09-252-991A-5717/c
; Sequence 5717, Application US/09252991A
; Patent No. 6551795
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-961-083-207
;
; Alignment Scores:
; Pred. No.: 230 Length: 805
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.60% Indels: 0
; DB: 3 Gaps: 0
;
; US-10-066-521-6 (1-1344) x US-08-961-083-207 (1-805)
;
; Qy 918 LeuThrHisLeuSerLeuSerMet 925
; Db 801 TTAACCCACTTATCATTCATG 778
;
; RESULT 7
; US-036-784-207/c
; Sequence 207, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 207:
;
; US-09-536-784-207
;
; Alignment Scores:
; Pred. No.: 230 Length: 805
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.60% Indels: 0
; DB: 4 Gaps: 0
;
; US-10-066-521-6 (1-1344) x US-09-536-784-207 (1-805)
;
; Qy 918 LeuThrHisLeuSerLeuSerMet 925
; Db 801 TTAACCCACTTATCATTCATG 778
;
; RESULT 8
; US-09-489-039A-2718
; Sequence 2718, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2718
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
;
; US-09-489-039A-2718
;
; Alignment Scores:
; Pred. No.: 244 Length: 855
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.60% Indels: 0
; DB: 4 Gaps: 0
;
; US-10-066-521-6 (1-1344) x US-09-489-039A-2718 (1-855)
;
; Qy 830 GlnCysAlaLeuGlnLysLeuIle 837
; Db 689 CAGTGGCGCTTTACAAAACTTATC 712
;
; RESULT 9
; US-09-252-991A-5722
; Sequence 5722, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; CORRESPONDENCE ADDRESS:
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5722
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5722

Alignment Scores:
Pred. No.: 254 Length: 891
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-252-991A-5722 (1-891)

Qy 849 GlnSerLeuAlaSerAlaLeuVal 856
Db 316 CAGAGCCTGGCATGGCGCTGGTG 339

RESULT 10
US-09-262-856A-13/c
; Sequence 13, Application US/09262856A
; Patent No. 6333164
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, Kazutosh
; APPLICANT: MIZUTANI, Shigetoshi
; APPLICANT: ENDO, Masahiro
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1422-372P
; CURRENT APPLICATION NUMBER: US/09/262,856A
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-262-856A-13

Alignment Scores:
Pred. No.: 269 Length: 944
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-262-856A-13 (1-944)

Qy 797 IleLeuThrSerProSerLeu 804
Db 713 ATCTTGACCACTTCACCATCTT 690

RESULT 11
US-08-632-514C-9/c
; Sequence 9, Application US/08632514C
; Patent No. 5834234
; GENERAL INFORMATION:
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,177
; FILING DATE: 2
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
```

```
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/632,514C
; FILING DATE: 29-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.188
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8459
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-632-514C-9

Alignment Scores:
Pred. No.: 273 Length: 958
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-632-514C-9 (1-958)

Qy 1218 LeuGlyArgGluLeuSerSerArg 1225
Db 583 CTTGGGCGAGAGCTGTCATCCCG 560

RESULT 12
US-09-188-177-9/c
; Sequence 9, Application US/09188177
; Patent No. 6057132
; GENERAL INFORMATION:
; APPLICANT: GALLO, Gregory J.
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,177
; FILING DATE: 2
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
```



```

; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.188
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8459
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-188-177-9
Alignment Scores:
Pred. No.: 273 Length: 958
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 3 Gaps: 0

US-10-066-521-6 (1-1344) x US-09-188-177-9 (1-958)
Qy 1218 LeuGlyArgGluSerSerArg 1225
Db 583 CTTGGCGAGAGCTGTCATCCGG 560

RESULT 13
US-08-858-207A-40/C
; Sequence 40, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

US-08-858-207A-40
Alignment Scores:
Pred. No.: 319 Length: 1117
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-858-207A-40 (1-1117)
Qy 918 LeuThrHisLeuSerLeuSerMet 925
Db 192 TTAACCCACTATCATTCATG 169

RESULT 14
US-08-961-527-13/C
; Sequence 13, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-13
Alignment Scores:
Pred. No.: 321 Length: 1126
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

US-10-066-521-6 (1-1344) x US-08-961-527-13 (1-1126)
Qy 775 LeuLeuGluSerLeuArgLeuAsp 782
Db 1088 CTGCTGGAGTCACTTCGACTGGAT 1065

RESULT 15

```

```

US-09-491-577-75/c
; Sequence 75, Application US/09491577
; Patent No. 6610511
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Carlson, John R.
; APPLICANT: Kim, Hunhyong
; APPLICANT: Clyne, Peter J.
; APPLICANT: Warr, Coral G.
; TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila
; FILE REFERENCE: 44574-5061-US
; CURRENT APPLICATION NUMBER: US/09/491,577
; CURRENT FILING DATE: 2000-01-25
; EARLIER APPLICATION NUMBER: US 60/117,132
; EARLIER FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
; OTHER INFORMATION: DORLU 9.1
US-09-491-577-75

```

```

Alignment Scores:
Pred. No.: 336 Length: 1176
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.60% Indels: 0
DB: 4 Gaps: 0

```

US-10-066-521-6 (1-1344) x US-09-491-577-75 (1-1176)

```

Qy 34 LysLysLysSerSerGluSerThr 41
Db 644 AAAAAGAGAGCGAGCGAGTCCACG 621

```

Search completed: July 16, 2004, 18:08:22
Job time : 209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 13:36:00 ; Search time 1299 Seconds
(without alignments)
5047.629 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: Oligo

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3190992 seqs, 2439311697 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6373542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10066521/runat_13072004_122319_11640/app_query.fasta_1.1543
-DB=Published_Applications_NA -Qfmt=fastap -SUFFIX=Oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10066521@cgn2_1.1.876 @runat_13072004_122319_11640
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

RESULT 1
US-10-124-498-5
; Sequence 5, Application US/10124498
; Publication No. US20030017983A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-367001
; CURRENT APPLICATION NUMBER: US/10/124,498
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 10/066,521
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4035

ALIGNMENTS

1	1344	100.0	4035	13	US-10-124-498-5	Sequence 5, Appli
2	1344	100.0	4035	15	US-10-066-521-5	Sequence 5, Appli
3	1006	74.9	3489	13	US-10-416-642-3	Sequence 3, Appli
4	1006	74.9	3926	15	US-10-216-645-1	Sequence 1, Appli
5	988	73.5	3830	15	US-10-216-645-3	Sequence 3, Appli
6	968	72.0	3226	13	US-10-092-900A-347	Sequence 347, App
7	616	45.8	3900	12	US-10-677-943-23	Sequence 23, Appl
8	616	45.8	3900	13	US-10-399-443-23	Sequence 23, Appl
9	385	28.6	1157	12	US-10-677-943-1	Sequence 1, Appli
10	385	28.6	1157	13	US-10-399-443-1	Sequence 1, Appli
11	196	14.6	1075	12	US-10-677-943-3	Sequence 3, Appli
12	196	14.6	1075	13	US-10-399-443-3	Sequence 3, Appli
13	57	4.2	2099	13	US-10-027-632-258159	Sequence 258159,
14	57	4.2	2099	16	US-10-027-632-258159	Sequence 258159,
15	12	0.9	509	15	US-10-029-386-10957	Sequence 10957, A
16	12	0.9	1704	15	US-10-029-386-24660	Sequence 24660, A
17	12	0.9	3108	10	US-09-965-621-23	Sequence 23, Appl
18	12	0.9	3108	16	US-10-407-866-23	Sequence 23, Appl
19	12	0.9	3186	13	US-10-124-498-17	Sequence 17, Appl
20	12	0.9	3186	15	US-10-066-521-17	Sequence 17, Appl
21	12	0.9	3218	16	US-10-407-866-67	Sequence 67, Appl
22	12	0.9	3447	12	US-10-677-943-5	Sequence 5, Appli
23	12	0.9	3447	13	US-10-399-443-5	Sequence 5, Appli
24	12	0.9	3466	16	US-10-108-260A-718	Sequence 718, App
25	12	0.9	4931	15	US-10-028-374-1	Sequence 1, Appli
26	12	0.9	4931	15	US-10-183-770-1	Sequence 1, Appli
27	10	0.7	610	13	US-10-424-599-139217	Sequence 139217,
28	9	0.7	240	15	US-10-029-386-16506	Sequence 16506, A
29	9	0.7	270	12	US-09-922-293-1943	Sequence 1943, Ap
30	9	0.7	339	17	US-10-437-963-46914	Sequence 46914, A
31	9	0.7	440	13	US-10-424-599-35497	Sequence 35497, A
32	9	0.7	564	15	US-10-029-386-2806	Sequence 2806, Ap
33	9	0.7	578	10	US-09-764-891-1745	Sequence 1745, Ap
34	9	0.7	720	13	US-10-027-632-150437	Sequence 150437,
35	9	0.7	1026	16	US-10-027-632-150437	Sequence 150437,
36	9	0.7	1033	17	US-10-437-963-16592	Sequence 16592, A
37	9	0.7	1033	13	US-10-412-699B-1129	Sequence 1129, Ap
38	9	0.7	1033	16	US-10-374-780A-676	Sequence 676, App
39	9	0.7	1296	13	US-10-412-699B-1128	Sequence 1128, Ap
40	9	0.7	1296	16	US-10-374-780A-675	Sequence 675, App
41	9	0.7	1842	13	US-10-424-599-80385	Sequence 80385, A
42	9	0.7	1907	13	US-10-027-632-100254	Sequence 100254,
43	9	0.7	1907	16	US-10-027-632-100254	Sequence 100254,
44	9	0.7	2596	13	US-10-653-595-99	Sequence 99, Appl
45	9	0.7	2596	13	US-09-397-945-99	Sequence 99, Appl

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4032)
; US-10-124-498-5

Alignment Scores:
Pred. No.: 0 Length: 4035
Score: 1344.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-124-498-5 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAGGAAGAAATTTACAGACATTCAGGAATTTACTAAAGAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATTCAGAAATGCAACGTTGGAATGTCGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIlePheGlu 80
Db 181 CTCITGCATGAGTATTATGGAGCATCGCTGGCTGGGCTACGTCCATTAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCTCTCGAGAAAGGACCGGATGACATGAAAAAAATTTCAAA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlnIleSerGlnAla 120
Db 301 GCTATGGAAACAAAGAGGTGCGACAGCAGCAGACAGAAAGAAATTTCAAGCT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlyHisGlyAspThr 140
Db 361 ATGGAACAAAGAGGTGCGACAGCAGCAGACAGAAAGAAAGCAGATGGAGGTGACACA 420
Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGGACTACAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTACGTCGTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAAACATGCTGCTGACTGGCCGGAATGCAACGTTGGCTTTTGATTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys 200
Db 541 GACCGGTGGGCTTCGGGCTCGCACGCTGCTCTGACGAAAGTCAAGAAATTTGGGAAA 600
Qy 201 SerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMet 220
Db 601 TCGGCTCTAGCCAGGAAGATCGTGTGCTGGCGCGCAAGGTGACCTTACACAGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTCCTCTACGCTTCTTCCTCCCGTTTAGAGAGATGACAGCGGAAGAGAGAGAGAGTGTC 720
Qy 241 ThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSer 260
Db 721 ACAGAGTTCACTCCAGGGAGTGGCCAGACTCCACAGGCTCCGGTGGACGGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAGAAAGGCTGTTGTTTCATCATTCACGCTTTCGATGACCTGGGCTCTGCTCTCAAC 840
```

```
Qy 281 AsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIleArg 300
Db 841 AATGACACAAAGCTCTGCAAGAGCTGGGCTGGAAGCAGCGCTCGTTTCCACCCCTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCTCTGCTCCCTGAGTCTCTCTGATCGTCAACGTCAGACGCTG 960
Qy 321 GlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIle 340
Db 961 GGCACAGAGAAGCTCAAGTCAGAGGTGCTGCTCCCGCTTACCTGTTAGTAGGAATC 1020
Qy 341 SerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLysThr 360
Db 1021 TCGGGGAAACAAAGAAATCCACTTCTTGGCGGGGATTTGGTGAGCAATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGGTTGCTGGCATCATGAACACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCC 1140
Qy 381 ValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCTGGCGCTGCACTGCAGACGCTGGTGGGGGAGAGCGTCGCC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACGCTCACAGGCTGCGCGCTTTTGTGTTTCATCAGCTCACCCCT 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGCTGCTCGCGCTGCTCAATCTGGAGGAAGAGTGTCTCTGAAGCGCTCTGTC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet 460
Db 1321 CGTATGCTGTGGAGGAGTGTGGAATAGGAAGTCACTGTTTGTGATGTCAGCACCTCATG 1380
Qy 461 ValGlnGlyLeuGlyLysSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGAGTCTGAGCTCCGCTGCTGTTTCCATGATGAACATCTCTTCTCCCA 1440
Qy 481 AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACTTCTCCACTCAGTCTCCAGGAGCTTCTGTGCC 1500
Qy 501 AlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal 520
Db 1501 GCCTTGTTACTACTGTTAGAGGCTTGAATTCGAGCCAGCTCTCTGCCCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAAGCAAAAGAGTCCATGGAGCTTAAACAGCGAGGCTTCCATATCCACTCGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu 560
Db 1621 ATGAAGCGTTTCTTTGTCCTCGTGAGCGAAGAGCTAAGAGGAGCCACTGGAGGTCTCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCTTCCCTGGGGGTGAAGCAGAAAGCTTCTGCACCTGGGTCTCTCTGTTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGCAGGCTTAATGCCACCCAGGAGACACCTTGAGCCCTTCCACTGTCTTCTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCAAAGAGGTGGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACAGAACCTGGAGCTTGTATGACATCTTCTTCTGCTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAla 660
```

1921 TTCCGGAATAATTCGGGTGGATGTCAAAGGGATCTTTCCAAAGAGATGAGTCCGCTGAGCA 1980
661 CysProValProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAsp 680
1981 TGTCTCTGTGGTCCCTCTATGGATGCGGATAGAGACCTTCATTGAGAGCAGTGGGAAGT 2040
681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
2041 TTCTGCTCCATGCTTGGCACCACCCACACACCTTCGGCGAGCTGGACCTGGCGCAGCAGCATC 2100
701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysAlaIle 720
2101 CTGACAGACGGCGGCATGAAGACCTGTGTGCCAAGCTTGAGGCATCCCACTGCAAGATA 2160
721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
2161 CAGACCTGTATGTTAGAAATGACAGATTACCTCTGGTGTGCAGCACCTCTGGAGAAATC 2220
741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
2221 GTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGGAAGAG 2280
761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
2281 GATGTAAAGGATGGCGTGTGAAGCCTTAAACACACCCAAATGTTGTGTGGAGTCTTTGAGG 2340
781 LeuAspCysCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThrThr 800
2341 CTGGATTGCTGTGGATTGACCCATGCCCTGTACCTGGAAGATCTCCCAATCCTTACGACC 2400
801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
2401 TCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAGGGAGTAATG 2460
821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
2461 CCTCTCAGTGTGCTTCCAGAGTCTCCAGTGCGCCCTGCAGAGAGCTGATATCGAGGAGC 2520
841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
2521 TGTGGCATCACAGCCAGCGGTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGC 2580
861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLysCys 880
2581 TTGACACACCTGTGCTATCCAAACACAGCCTGGGGAGCAAGGTGTAAATCTACTGTGT 2640
881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGGCCACCTG 2700
901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaMetGlyAsnSerTrpLeuThrHis 920
2701 GACAGCGCTGGCTGTGGTTTCTTGCACTTGGCCTTATGGGTAACTCAATGCTGACGCAC 2760
921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLysCysGluValMet 940
2761 CTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGGAGGTCAATG 2820
941 ArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAla 960
2821 AGAGAACAATCTTGTCACTCCAGGACCTGGAGTTGGTAAAGTGTCACTCACGCCCGCG 2880
961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
2881 TGCTGTGAGAGTCTGTCTGTGTGATCTCGAGAGCAGACACCTGAAGAGCCTGATATCTC 2940
981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLysGlnLys 1000
2941 ACGGACAAATGCCCTGGGTGACGGTGGGTGTGCTGCTGTGCGAGGAGCTGAGCAAAAG 3000
1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGlu 1020

3001 AACAGTGTCTTCAGACGAGACTCGGTTCAAGGCATGTGAGACTGACTTCTGATGTGTGTGAG 3060
1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
3061 GCATCTCTCTTGGCCCTTTCCTGCAACCGGCATCTGACAGTCTTAAACCTGTGTCAGNAT 3120
1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTCCTGCTGTCCCACGCTCAAC 3180
1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeuGlu 1080
3181 TTACAGATAAATTTGGGCTGTGGAATGCGAGTACCTCTGTGCAATAAGGAAGCTCTGGAG 3240
1081 GluValGlnLeuLysProArgValIleAspGlySerTrpHisSerPheAspGlu 1100
3241 GAAGTGAGACTACTCAAGCCCGAGTCTGAATGACCGGTAGTTGGCATTCTTTTGTATGAA 3300
1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
3301 GATGACCGACACAAATAGGACTTACTTTCCGGCTCCCTGAAAGCCGGCATGGCCATGT 3360
1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysIleArgValSerLeuLeuAlaGlyAsp 1140
3361 GCCTTGTCTGGGGGATGAACCCAGACAGAGAAGAGCGTGTGTCGCTTCTGGCTGGAGAC 3420
1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
3421 TTCAGAGCAGTACACCATTTGCCAAGTCTCTGCTTGGCCACGGCAATATGGTAGTCC 3480
1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
3481 CAGAGAGTTGCACACAGTGGAGCAGAGCTCCCGCAACCCATGGCAGGCAGCAACACAAA 3540
1181 GlnAspLysMetLeuSerValGlyTrpSerGlyAlaTrpSerGluThrAlaGluLeuGlu 1200
3541 CAAGATAAAATGTTGAGTGTGGATATTCGGAGCCTGGTCTGAACCTGCTGAGCTCGAA 3600
1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyMetAlaTrpSerLeuGlyArg 1220
3601 GGGCTTGATCCAACTGCTGATCATGACACGAGGATATGGCTTGGTCTACTAGGAGGA 3660
1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
3661 GAGCTGAGCTCGAGGGCTTGTGTCCAAAGTGTCTGATGACCAACAGCGGTGTGTCTTGGT 3720
1241 HisTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
3721 CACTGGAGCGGCTGGGCTCTAGGGGCTGGTGTCTTAACAGTGTCTGATGACCAACAGCGGT 3780
1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
3781 GTGTCTTGTTCATCTGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCAACAGTGTCTGATGAC 3840
1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
3841 CACAGCGTGTGGCTTGTGTCTCACTGGAGCGCTGGGCTCGAGGGGCTTGTGTCTCAACAGT 3900
1301 AlaAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuVal 1320
3901 GCTGATACCAACAGCGGTGTGTCTTGGTCACTGGAGCGGCTGGGCTCGAGGGGCTTGTG 3960
1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
3961 TCCAAACAGTGTGATGACCAACAGCGGTGTGTCTTGGTCACTGGAGAGCGGCTGGGCTCGAG 4020
1341 GlyLeuValSer 1344
4021 GGGCTGGTGTCT 4032

RESULT 2

US-10-066-521-5

; Sequence 5, Application US/10066521

; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334001
; CURRENT APPLICATION NUMBER: US/10/066,521
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4032)
US-10-066-521-5

Alignment Scores:
Pred. No.: 0 Length: 4035
Score: 1344.00 Matches: 1344
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-066-521-5 (1-4035)

Qy	1	MetGluGlyAspLysSerLeuThrPheSerTyrGlyLeuGlnTyrCysLeuTyrGlu	20
Db	1	ATGGAGAGAGCAATCGCTCACCTTTCCAGCTACGGCTGCAATGGTGTCTATGAG	60
Qy	21	LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerGluSer	40
Db	61	CTAGACAGGAGGAATTTACAGACATTCAGGAATTTACTAAAGAGAAATCTTCAGAAATCG	120
Qy	41	ThrThrCysSerIleProGlnPheGluLeuGlnAlaLeuValGluCysLeuAlaLeu	60
Db	121	ACCCATGCTCTATTCCACAGTTTGAATTCAGAGATGCCACGTGGAATGTCGGCATC	180
Qy	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrSerIleSerIlePheGlu	80
Db	191	CTCTTGATGATGATATGAGGATCGCTGGCTGGGCTAGCTCCATAGCATCTTTGAA	240
Qy	81	AsnMetAsnLeuAthrThrLeuSerGluLysAlaArgAspMetLysLysIleSerGln	100
Db	241	AACATGAACCTCGCAACCTCTCGAGAGGACCGGATGACATGAAAAAATTTCAAA	300
Qy	101	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlnIleSerGlnAla	120
Db	301	GCTATGGAACCAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAAGAAATTTCAAGCT	360
Qy	121	MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGlyGlyAspThr	140
Db	361	ATGGACCAAGAGGTGCCACAGCAGCAGACAGACAGACAGACATGGAGGTGACACA	420
Qy	141	TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgSer	160
Db	421	TGGACTACAAAGAGTCAGCTGATGACCAAAATTCGCTGAGGAGGAGGATGACGTCGTAGT	480
Qy	161	PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer	180
Db	481	TTTGAAAACTGCTGCTGACTGGCCGGAATGCAACGTTGGCTGGTCTTTGATTCA	540
Qy	181	AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys	200

Db	541	GACCGTGGGGCTTCGGCCCTCGCACGGTGGTTCTGCACCGAAAGTCAGGAATGGGAAA	600
Qy	201	SerAlaLeuAlaArgArgIleValLeuCysTyrTrpAlaGlnGlyGlyLeuTyrGlnGlyMet	220
Db	601	TCGGCTTAGCCAGAGGATCGTGTGTGGCGCAAGGTGACTCTACACAGGAATG	660
Qy	221	PheSerTyrValPhePheLeuProValArgGluMetClnArgLysLysGluSerSerVal	240
Db	661	TTCTCTTACCTCTTCTTCTCCCGTTAGAGATGCGCGGAGAGAGAGACGCGTGC	720
Qy	241	ThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGluIleMetSer	260
Db	721	ACAGAGTTCACTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGGAGATCATGTCC	780
Qy	261	ArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsn	280
Db	781	CGACCAAGAAAGGCTGTTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGCTCAAC	840
Qy	281	AsnAspThrLysLysCysLysAspTyrAlaGluLysGlnProPheThrLeuIleArg	300
Db	841	AATGACACAAAGCTCTGCAAGACTGGGCTGGAAGCAGCCTCCGTTTCAACCTCATACGC	900
Qy	301	SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal	320
Db	901	AGTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCTTCTGATCGTCACCGTCAGACGCTG	960
Qy	321	GlyThrGluLysLysSerGluValValSerProArgTyrLeuLeuValArgGlyIle	340
Db	961	GGCACAGAGAGCTCAAGTCAGAGGTGTGTCTCCCGCTTACCTGTGTAGTAGGAATC	1020
Qy	341	SerGlyGluClnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThr	360
Db	1021	TCCGGGGAACAAAGAAATCCACTTGTCTTGAGCGGGGATTTGGTGAAGCATCAGAAGACA	1080
Qy	361	GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla	380
Db	1081	CAAGGTTGGTGGGATCATGAACACCGTGAGCTGCTCGACCAGTCCAGGTGCGCGCC	1140
Qy	381	ValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla	400
Db	1141	GTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGACGCTGGTGGGGAGAGCGTCCGC	1200
Qy	401	ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro	420
Db	1201	CCCTTCAACCAACGCTCACAGGCTGCAGCGCGCTTTGTGTTTCATAGCTCACCCCT	1260
Qy	421	ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys	440
Db	1261	CGAGGCGTGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGCTCTGAAGCGCTTCTGC	1320
Qy	441	ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMet	460
Db	1321	CGTATGGCTGTGGAGGAGTGTGGAATAGGAATGAGTGTGTTGATGTGTCAGCACCTCATG	1380
Qy	461	ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuPro	480
Db	1381	GTTCAAGGACTCGGGGAGTCTGAGCTCCGGTGTCTGTTTTCATGAACATCTCTTCTCCCA	1440
Qy	481	AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla	500
Db	1441	GACAGCCACTGTGAGGAGTACTACACTTCTCCACTCTCAGTCTCCAGGACTTCTGTGCC	1500
Qy	501	AlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal	520
Db	1501	GCCTTGCTACTACGTGTAGAGGCGCTGGAATCGAGCCAGCTCTCTGCCCTCTGTACGTT	1560
Qy	521	GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyr	540
Db	1561	GAGAGCAAGAGGTCCATGGAGCTTAAACAGCAGGCTTCCATATCCACTCGCTTTGG	1620
Qy	541	MetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluValLeu	560
Db	1621	ATGAAGCGTTTCTTGTGGCCTCGTGAGCGAAGACGTAAAGGAGGCCACTGGAGGTCCTG	1680

Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGAGCTTCCTGCACTGGGTCTCTCTGTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrTrpProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTACAGACCCCTAAATGCCACACCCAGAGAGACCCCTGGACGCCCTCCACTGTCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTGTTGGCTTGGCATTTAAACAGCTTCCAAGAAAGTGTGGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACAGAACTTGGACTTGTAGCATCTTCTGCTCCAGCACTGTCCGCTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAla 660
Db 1921 TTGCGGAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGTCCCTCTATGATGCGGATAGACCCCTCATTTGAGGAGCAGTGGGAAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTTCTGCTCCATGCTTGGCACCCACCCACCTGGGGCAGCTGGACCTGGGGCAGCAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 2101 CTGACAGACGGGGCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCCAACCTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCCCTGATGTTTAGAAATGACAGATTACCCCTGGTGTGCAGCACCTCTGGAGAATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTGGGGAGCACCACCTGCAAGGAAGAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
Db 2281 GATGTAAGGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGGATTCTGTGGATTGACCCATGCCCTGTTTACCCTGAAGATCTCCCAAAATCCTTACCAGC 2400
Qy 801 SerProSerLeuLysSerSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCACGCTGAAATCTCTGAGCTTGGCAGGAAACAGGTGACAGACCAGGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTCAGTGTGCTTCAGAGTCTCCAGTGCGCCCTGACAGAGCTGATATCTGGAGC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCCACGGGTGGCAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880
Db 2581 TTGACACACCTGTGCTATCCACACAGCCTGGGGACAGAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTCTAGTCTGCAGAGGCTGATGCTGAATCAGTGCACACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGGCTGGCTGGGTGTTTCTTGGCATCTTGGCTTATGGGTAACTCATGTGCTCAGCCAC 2760

Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCCCTGTGGAAGACAAATGCGCTGAAGCTTCTGTGCGAGGTCAATG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeuThrAlaAla 960
Db 2821 AGAACAACATCTTGTCTATCTCCAGGACCTGGAGTGGTAAAGTGTCTATCTCACCGCGCG 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGCTGTGAGAGTCTGTCTCTGTGTGATCTCGAGGAGCAGACACCTGGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAATGCCCTGGGTGACGGTGGGGTGTGCTGCGCTGTGCGAGGAGCTGAACCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGTGTAAGGCATGTGGACTGACTTCTGATTGCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCACCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTGGTGCAGAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGCGSCCTTTGCTCTGCCACGCTTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 3181 TTACAGATAAATGGGCTGTGGAATGGCAGTACCCCTGTGCAAAATAAGAAAGCTCTCGAG 3240
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAGGTGAGCTACTCAAGCCCCGAGTGTGAATTGACGGTAGTTGGCAATCTTTTGTATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCCGGCTCCCTGAAAGCCGGCATGSCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysIleArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTGTGTGGGGGATGAACCCAGAGCAGAAAGCGTGTGCGCTTCTGGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
Db 3421 TTCAAGAGCAGTACACGATTTGCCAAGTCTCTGCTGCGCCACGGCAATGGTGAATGCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTCACAAACGTGGAGCAGAGCTCCCGCNAACCCATGGCAGGCACGGAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTyrSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTAGTGTGGATATTCGGAGCCTGGTCTGAAAACCTGCTAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyClyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGGCTTGATCCCAACAGTGTGATCATGACCGAGGTATGGCTTGGTCACTAGGGAGA 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTGTCTGATGACCAAGCGGTGTGCTCTGGT 3720
Qy 1241 HisTrpGluArgLeuGlySerArgGlyTyrCysLeuAsnSerAlaAspHisSerGly 1260
Db 3721 CACTGGAGCGGCTGGGCTCTAGGGGCTGGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780
Qy 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyValSerHisSerAlaAsp 1280
Db 3781 GTGTCTGGTCACTGGAGCGGCTGGGCTCGAGGGCTTGTGTCCAAACAGTGTGATGAC 3840
Qy 1281 HisSerGlyValAlaTrpSerLeuGlyAlaAlaGlyLeuGlyLeuValSerAsnSer 1300

QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 1894 TCTCTGTTGGGTGAGCAGCTAATGCCACCAACCCAGGAGACACCCCTGGACGCTTCCAC 1953
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTCTGCTTGGCATATAACAGCTTCCAGAA 2013
QY 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 2014 GTGTGGCTTCGGATTAAACAGAACCTGGACTGTATAGCATCTTCTCTGCTCCAGCAG 2073
QY 638 CysProTrpLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGlnSer 657
DB 2074 TGTCCGTTATTTGCGGAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCC 2133
QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 2134 GCTGAGGATGTCCTGTGTGCTCTATGGATGCGGGATAAGACCTCATTTGAGAGCAG 2193
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 2194 TGGGAAGATTCTGCTCCATGCTTGGCACCCACCCACACCTGGCGCAGCTGGACCTGGGC 2253
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 2254 AGCAGCATCTTCACAGAGCGGGCCATGAAGCCCTGTGTGCCAAGCTGAGGCATCCAC 2313
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 2314 TGCAGATATACAGACCTCATGTTTGAATAATGCACAGATTACCCCTGGTGTGCAGCACCTC 2373
QY 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB 2374 TGGAGAACTGTCATGCGCCACCGTAACCTAAGATCCCTCAACTGGGAGGCCACCCACTG 2433
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB 2434 AAGGAAGAGGATGTAAGATGCGGTGTGAAGCCTTAAACACCCCAAAATGTTGTGGAG 2493
QY 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
DB 2494 TCTTTGAGCGTGATGCTGTGGATTGACCCATGCTGTATACCTGAAGATCTCCCAATC 2553
QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB 2554 CTTAGCAGCTCCCCAGCGCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACAGACAG 2613
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB 2614 GGAGTAATGCTCTCAGTGTATGCTTTGAGAGTCTCCAGTGGCGCTGCAGAAAGCTGATA 2673
QY 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB 2674 CTGGAGGACTGTGGCATACAGCCACCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGAGCTGTACACACTGTGCTCTATCCAAACACCGCTGGGGAACGAAGGTGTAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTGATCCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATGCTGATCATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCCACCTGGACACCGCTGGCTGTGTTTCTTGTGCACCTTGGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGACCTTGGACCTTAGCATGAACCTGTGGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957

DB 2974 GAGGTCAATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCGCCGGTGTGTGAGAGTCTGTCTCTGTGTATCTCGAGGAGCAGACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAGGACATGCCCTGGGTGACGGTGGGTGTCTGCACCTGTGCGAGGAGCTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGTACGAGACTCGGGTTGAAGGCATGTGGAAGTGTCTGAT 3213
QY 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
DB 3214 TGCTGTGAGGACACTCTCTTGGCCCTTCTCTGCACCCGACATCTGACAGTCTAAACCTG 3273
QY 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
DB 3274 GTGCAGATAAATTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCTTTGCTGTCTCCC 3333
QY 1058 ThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
DB 3334 ACGTCTAACTTACAGATAATTTGGCTGTGGAATGGCAGTACCTGTGCAAAATAAGGAAG 3393
QY 1078 LeuLeuGluValGlnLeuLeuLeuLeuLeuProArgValIleAspGlySerTrpHisSer 1097
DB 3394 CTGTGGAGGAAGTGCAGTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTTGGCATTCT 3453
QY 1098 PheAspGluAspAspArg 1103
DB 3454 TTTGATGAGATGACCGG 3471
RESULT 4
US-10-216-645-1
; Sequence 1, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAELE
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN WATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3926
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-216-645-1
Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 15 Gaps: 0
US-10-066-521-6 (1-1344) x US-10-216-645-1 (1-3926)
QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
DB 454 ATTTCAAGCTATGGAAACAAAGAGGTGCCACAGCAGCAGACAGAAACAAGAAATT 513
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137

Db 514 TCACAGAGTATGGAAACAGAGGTGCCACACAGACAGACAGACAGAAACAGAGCATGGA 573
Qy GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 574 GGTGACACATGGGACTACAGAGTCCAGTATGACCAAAATTCGCTGAGGAGGAGATGTA 633
Qy ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTTGAAACACATGCTGCTGACTGGCCGGAATAATGCAAAACGTTGGCTGGCT 693
Qy PheAspSerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 694 TTTGATTACAGACCGGTGGGCTTCGGGCTCGACGGTGGTTCTGCACGGAAATCAGGA 753
Qy IleGlyLysSerAlaLeuAlaArgAlaValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 754 ATTTGGGAAATCGGCTTAGCCAGAGGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 813
Qy GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 814 CAGGGAATGTTCTCTACGTCTTCTTCTCCCGTTAGAGAGATGCAGCGGAAGAAGGAG 873
Qy SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCAGGGCTCCGGTACCGGAG 933
Qy IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 934 ATCATGTCCCGACAGAAAGCTGTGTTTCATCATTCACGGTTTCGATGACCTGGGCTCT 993
Qy ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 994 GTCTCTCAACAATATGACAAAGCTCTGCAAAAGACTGGGCTGAGAAGACGCTCCGTTCAAC 1053
Qy LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1054 CTCATACGCAGTCTGCTGAGGAAGTCTGCTCCCTCGAGTCTTCTCGATCGTCAACGTC 1113
Qy ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1114 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGTCTGTCCTCCCGCTTACCTGTAGTT 1173
Qy ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1174 AGAGAATCTCCGGGAAACAAGAATCCACTGTCTCTTGAGCGCGGATTTGGTGAGCAT 1233
Qy GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1234 CAGAAGACACAAGGTTGCGTGGCATCATGAACAACCGTGAGCTGCTCGACCACTGCCAG 1293
Qy ValProAlaValGlySerLeuIleCysValAlaIleuGlnLeuGlnAspValValGlyGlu 397
Db 1294 GTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGCGAGCTGCAGGACGTGGTGGGAG 1353
Qy SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1354 AGCGTGGCCCCCTTCAACCAAAACGCTCACAGCGCTGCACGCGCTTTGTTGTTTCATCAG 1413
Qy LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
Db 1414 CTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGATTGCTCTGAG 1473
Qy ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCTCGCGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTTGACCGGTGAC 1533
Qy AsnLeuMetValGlnGlyLeuGlyCysLeuSerGluArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGGTTCAGGACTCGGGAGTCTGAGCTCCGCTGCTCTGTTTCATGAACATC 1593
Qy LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497

Db 1594 CTTTCTCCACAGACAGCCACTGTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1653
Qy PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTTCTGTGCCCTTGTACTACGTGTGTAGAGGGCTCGAAATCGAGGCAGCTCTGTGCCCT 1713
Qy LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTAGCTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCCATATCCAC 1773
Qy SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 1774 TCGCTTTGGATGAAGCGTTTCTTGTGTGGCTCGTGAGCGAAGACGTTAAGAGGCCACTG 1833
Qy GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGGTCTGTGTGGCTGTCCCGTTCCCTGGGGGTGAAAGCAGAGCTTCTGTGCATCGGGTC 1893
Qy SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTGGGTGAGAGCCCTAATGCCACCCAGGAGACCCCTGGAGCGCTTCCAC 1953
Qy CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAGAA 2013
Qy ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGGCTTCGGATTAAACAGAACCTGGACTTGATAGCATCTTCTTCTGTCCCTCCAGCAC 2073
Qy CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCCGTATTTCGGAAAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC 2133
Qy AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
Db 2134 GCTGAGGCATGTCTGTGTGTCCTCTATGATGCGGGATAAGACCTTCATTGAGGAGCAG 2193
Qy TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTCTGCTCCATCTGTGACCCACCCACACCTGCGGCGAGCTGGACCTGGGC 2253
Qy SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTTGACAGAGCGGCCCATGAGACCTGTGTGCCAAGCTGAGGCGATCCACCC 2313
Qy CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAAGATACAGACCTGTGATGTTTAGAAATGCACAGATTACCCCTGCTGTGTCAGCACCTC 2373
Qy TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGAGAATCGTCAATGSCCAACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACCTG 2433
Qy LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2434 AAGAAGAGATGTAGGATGGCGTGTGAAGCTTAAACACCCCAAAATGTTTGTGGAG 2493
Qy SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCITTTGAGGCTGGATTGCTGTGTGATTGACCCATGCTGTTTACCTGAAGATCTCCCAAATC 2553
Qy LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTACGACCTTCCCGCAGCTGAAATCTCTGAGCGCTGGCAGGAAACAGGTTGACAGACCAG 2613
Qy GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAATATGCTCTCAGTGATGCTTGAGAGTCTCCAGTGCAGCTTGCAGAGAGCTGATA 2673
Qy LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGAGGAGCTGTGGCATCACAGCACGGGTTGCCAGAGTCTGGCCCTCAGCCCTCGTCAGC 2733

```
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGGAGCTTGACACACCTGTGCTATATCCAAACACAGCTGGGGAACAGGTGTAAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTGATCATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCCACCTGGACACGGCTGGCTGTGGTGTCTTCTGACCTTGCGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTGACGCACCTGAGCCTTAGCATGAACCTGTGGGAAGACATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
DB 2974 GAGGTTCATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValLysLeuSerArgSerArgHisLeuLysSer 977
DB 3034 ACCGCGCGGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACAGACCTGGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCACGACAATGCTGCTGACGCTGGGTGTGCTGCATCTGTGCGAGGAGCTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGCACTTCTGAT 3213
QY 1018 CysGlyGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
DB 3214 TGCTGTGAGGCACTCTCTCTGGCCCTTTCTGCAACCGGCACTGTACCAAGTCTAAACCTG 3273
QY 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
DB 3274 GTGCAAGATACTTCAGTCCCAAGAGATGATGAAGCTGTGTTCGGCCCTTTGGCTGTCCC 3333
QY 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTrpValGlnIleArgLys 1077
DB 3334 ACCTCTAACTTACAGATAAATGGGCTGTGGAAATGGCAGTACCCCTGTGCAAAATAAGGAAG 3393
QY 1078 LeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
DB 3394 CTGCTGGAGGAAGTGGAGCTACTCAAGCCCGGAGTGTGTAATTTGACGGTAGTTGGCATTTCT 3453
QY 1098 PheAspGluAspAspArg 1103
DB 3454 TTTGATGAAGATGACCGG 3471
RESULT 5
US-10-216-645-3
; Sequence 3, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAEL
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN WATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3830
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-10-216-645-3
Alignment Scores:
Pred. No.: 0 Length: 3830
Score: 988.00 Matches: 988
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.51% Indels: 0
DB: 15 Gaps: 0
US-10-066-521-6 (1-1344) x US-10-216-645-3 (1-3830)
QY 116 GluIleSerGlnAlaMetGluGlnGluGlyValAlaThrAlaAlaGluThrGluGluGlnGly 135
DB 451 GAAATTTTCAACAGCTATGGAACAAGAGGTGCCACAGCAGCAGCAGCAGCAAGAAGCAAGGA 510
QY 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
DB 511 CATGGAGGTGACACATGGGACTACAGAGTTCAGTGTGATGACCAAAATTCGCTGAGGAGGAG 570
QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
DB 571 GATGTAGCTGTGTAGTTTTTGAACAACATGCTGTGACTGGCCGGAAATGCAAAACGTTGGCT 630
QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
DB 631 GGTGCTTTTGTATTTCAGACCGGTGGGGCTTCGGGCTCGCACGGTGGTTCCTGCACGGAAG 690
QY 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
DB 691 TCAGGAATTCGGAAATCGGCTCTAGCCAGAGGATCGTGTGTGCTGGGCGCAAGGTGGA 750
QY 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
DB 751 CTCCTACCGAGGAATGTTCTCTCTAGCTTTCTCTCCCGTTAGAGAGATGCACGGGAAG 810
QY 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
DB 811 AAGGAGAGCAGTGTACAGAGATTTCATCTCCAGGGAGTGGCAGACTCCAGGCTCCGGTG 870
QY 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeu 275
DB 871 ACGGAGATCATGTCCCGCAGAAAGCTGTGTTCATCATTTGACGGTTCGATGACCTG 930
QY 276 GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
DB 931 GGCTCTGTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAGCCTCCG 990
QY 296 PheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
DB 991 TTCACCTCATACGACAGTCTGCTGAGGAAGTCTCTCTCTGAGTCTCTCTCATCGTC 1050
QY 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrIleu 335
DB 1051 ACCGTCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTGCTGTCTCTCCCGTTACCTG 1110
QY 336 LeuValArgGlyIleSerGlyGluGlnArgTleHisLeuLeuLeuGluArgGlyIleGly 355
DB 1111 TTGTTAGAGGAATCTCCGGGGAACAAGAATTCACCTTGCTCTTGAGCGCGGAGTTGGT 1170
QY 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
DB 1171 GAGCATCAGAAGACACAAGGGTTGCGTGGCATCATGAACACCGTAGCTGCTCGACCAG 1230
QY 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
DB 1231 TGCCAGGTGCGCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGCTGCAGAGCTGGTG 1290
QY 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
DB 1291 GGGGAGAGCGTGCCTCCCTTCAACCAACCGCTCAAGGCTTCACAGGCTTCGACCGCGCTTTTGTGTTT 1350
```

QY 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAenLeuGluAtrqValVal 435
DB 1351 CATAGCTCACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTGTCT 1410
QY 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
DB 1411 CTGAAGCGCTTCGCGGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAAGTGTTCAC 1470
QY 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
DB 1471 GGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTACATG 1530
QY 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeu 495
DB 1531 AACATCTCTTCCACAGACGACCTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTC 1590
QY 496 GlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeu 515
DB 1591 CAGGACTTCTGTCCGCTTGTACTACGTGTAGAGGCTTGGAAATCGAGCCAGCTCTC 1650
QY 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
DB 1651 TGCCCTCTGTAGTGTAGAGACAAAGAGTCCATGGAGCTTAAACAGGACAGGCTTCCAT 1710
QY 536 IleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
DB 1711 ATCCACTCGCTTGGATGAAGCGTTTCTGTGTGGCTCGTGAGCGAAGAGCTAAGGAGG 1770
QY 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
DB 1771 CCAGTGGAGGCTCTGGCTGCTCCCGTTCCTCCCTGGGGGTGAAGAGAGCTTCTGCAC 1830
QY 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
DB 1831 TGGGTCTCTGTGTGGGTGACAGCTAAATCCACCCAGGAGACACCCCTGGAGGCC 1890
QY 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
DB 1891 TTCACCTGTCTTTTCGAGACTCAGACAAAGAGTTTGTTCCTTGGCATTTAAACAGCTTC 1950
QY 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
DB 1951 CAAGAAGTGTGGCTCCGATTAACAGAACCTGGACTTGATAGCATCTTCCTCTGCGCTC 2010
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
DB 2011 CAGCACTGTCCGTATTTTCGGGAAATTCGGGTGGATGTCAAGGGATCTTTCCCAAGAGAT 2070
QY 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
DB 2071 GAGTCCGCTGAGGCATGTCTGTGCTCTATGGATGCGGGATAGAACCTCAITTGAG 2130
QY 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
DB 2131 GAGCAGTGGGAAGATTTCTGCTCCATGCTTGGCACCCACCCACACCTGGGCGACCTGGAC 2190
QY 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
DB 2191 CTGGGACAGCAGCATCTCTGACAGCGGGCCATGAAGACCTGTGTGTGCCAAGCTGAGCAT 2250
QY 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
DB 2251 CCCACCTGCAAGATACAGACCTCATGTTTGAAGATGCACAGATTACCCCTGGTGTGCAG 2310
QY 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThr 755
DB 2311 CACCTCTGAGAAATCGTATGGCCAAACCGTAACTAAGATCCCTCAACTGGGAGGCACC 2370
QY 756 HisLeuLysGluAlaAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
DB 2371 CACCTGAAGGAGAGGATGTGAAGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTTG 2430
QY 776 LeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795

DB 2431 TTGAGAGTCTTTGAGGCTGGATTCCTGTGGATTGACCCATGCCCTGTACCTGAAGATCTCC 2490
QY 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815
DB 2491 CAAATCTTTACGACCTCCCCACGCTGAAATCTCTGAGCCCTGGCAGAAACAAGGTGACA 2550
QY 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
DB 2551 GACCAGGAGTAATGCTCTCTCAGTGATGCTTCTGAGAGTCTCCAGTGCCTGCGCAAG 2610
QY 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
DB 2611 CTGATCTGGAGGACTGTGGCATCACAGCCACGGGTTGCCAGAGTCTGGCCCTCAGCCCTC 2670
QY 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
DB 2671 GTCCAGCAACCGAGCTTGACACACTGTGCTATCCACACAGACCTGGGGAACGAGGT 2730
QY 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
DB 2731 GTAAATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTG 2790
QY 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915
DB 2791 AATCAGTGCCACCTGACAGCGCTGGCTGTGGTTCCTTTCGACTTGGCTTATGGGGTAAC 2850
QY 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
DB 2851 TCATGGCTGACGACCTGAGCCTTAGCATGAACCTGTGGAGACAATGGCGTGAAGCTT 2910
QY 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
DB 2911 CTGTGCGAGTCTATGAGAGAACCATCTTGTCACTCCAGGACCTGGAGTTGGTAAAGTGT 2970
QY 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975
DB 2971 CATCTCACCCCGCTGTGTGAGAGTCTGTCTCTGTGATCTCGAGGACACACACCTG 3030
QY 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
DB 3031 AAGAGCTGTGATCTCACGGACATGCTCCCTGGGTGACGGTGGGTTGCTGCACCTGTGCGAG 3090
QY 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
DB 3091 GGAAGTGAAGCAAAAGAACAGTGTCTGACAGACTCGGTTGAAGGCATGTGGACTGACT 3150
QY 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
DB 3151 TCTGATTGTGTGAGGACCTCTCTTGGCCCTTCTGCAACCGGCATCTGACACAGTCTA 3210
QY 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
DB 3211 AACCTGGTGAGAAATAACTTCACTCCCAAGGAATGATGAAGCTGTGTTTCGGCTTTGCC 3270
QY 1056 CysProThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
DB 3271 TGTGCCACGCTCAACTTACAGATAAATGGGCTGTGAAATGGCAGTACCTCTGTGCAATA 3330
QY 1076 ArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValIleAspGlySerTrp 1095
DB 3331 AGGAAGCTGCTGAGGAGAGTGCAGCTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTTGG 3390
QY 1096 HisSerPheAspGluAspArgArg 1103
DB 3391 CATCTTTTGTGAAGATGACCGG 3414

RESULT 6

US-10-092-900A-347
; Sequence 347, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara

```

; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 347
; LENGTH: 3226
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)..(3168)
US-10-092-900A-347

Alignment Scores:
Pred. No.: 0
Score: 968.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 72.02%
DB: 13
Length: 3226
Matches: 968
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

```

US-10-066-521-6 (1-1344) x US-10-092-900A-347 (1-3226)
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
Db 249 CATGGAGTGCACATGGGACTACAGAGTCACGTGATGACCAATTCGTCTGAGGAGGAG 308
Qy 156 AspValArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 309 GATGTACGTCTAGTATTTTGAATAACACTGCTGCTGACTGGCGGAAATGCAAAACGTTGGCT 368
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValLeuHisGlyLys 195
Db 369 GGTGCTTTTGAATTCAGACCGGTGGGGCTTCGGCTCCGACGGTGGTTCGCGGAAAG 428
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 215
Db 429 TCAGGAATTCGGAAATCGGCTCTAGCCAGGAAGATCGTGTCTGTCTGGGCGCAAGGTGGA 488
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 489 CTCTACCCAGGGAATGTTCTCTACGTTCTCTCTCCCGTTAGAGAGATGTCAGCGGAAG 548
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 549 AAGGAGAGAGTGTCTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTG 608
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
Db 609 ACGGAGATCATGTCCCGACCAAGAAAGGCTGTTGTTTCATCATTTGACGGTTTCGATGACTG 668
Qy 276 GlySerValLeuAenAenAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
Db 669 GGCTCTGTCTCAACATGATGACAAAGCTCTGCAAGAGCTGGGCTGAGAAAGCAGCCTCCG 728
Qy 296 PheThrLeuIleArgSerLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
Db 729 TTCACCTCATACGAGTCTGCTGAGGAAGTCTCTCTCTCTGAGTCTCTCTCTGATCGTC 788
Qy 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
Db 789 ACCGTACAGACGCTGGGCACAGAGAGTCTCAAGTCAGAGGTCGTGTCCTCCCGTTACTG 848
Qy 336 LeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGly 355
Db 849 TTAGTTAGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTCTTGAAGCGGAGTTGGT 908
Qy 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
Db 909 GAGCATCAGAAAGACACAAAGGTTGCGTGCATCATGAACCAACCGTGAGCTGCTCGACCAG 968
Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db 969 TGCAGGTGCCCGCGTGGGCTCTCTCATCTGCTGGTGGCCCTGACGCTGCGAGGAGTGGT 1028
Qy 396 GlyCysSerValAlaProPheAenGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
Db 1029 GGGGAGAGCGTCGCCCTTCAACCAACCGCTCACAGGCTGACAGCGCTTTTGTGTTT 1088
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAenLeuGluAlaGluValVal 435
Db 1089 CATCAGCTCACCCCTCGAGGCGTGGTCCGGCGTGTCTCAATCTGGAGGAAGAGTTGTC 1148
Qy 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
Db 1149 CTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGGGAATAGGAAGTCAAGTGTGTTGAC 1208
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
Db 1209 GGTGACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCGGTGCTCTGTTTTCATG 1268
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeu 495
Db 1269 AACATCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1328

```

QY 496 GlnAspPheCysAlaAlaLeuTyrTyrValLeuGlnGlyLeuGluIleGluProAlaLeu 515
Db 1329 CAGGACTTCTGTGGCGCTTGTACTACGTGTGTAGAGGGCCCTGGAAATCGAGCCAGCTCTC 1388
QY 516 CysProLeuTyrValGluIleThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db 1389 TGCCCTCTGTACGTTGAGAACAAAGAGGTCCTATGGAGCTTAAACAGGCGAGGCTTCCAT 1448
QY 536 IleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
Db 1449 ATCCACTCGCTTGGATGAAGCGTTTCTTGTGGCTCGTGAGCGAAGACGTAAGGAGG 1508
QY 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
Db 1509 CCACCTGGAGGTCCTGTGGGCTGTCCCGTTCCTCCCTGGGGGTGAAGCAGAAAGCTTCTGCAC 1568
QY 576 TrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
Db 1569 TGGGTCTCTGTGGGTGACAGGCTTAATGCCACCCCGAGAGACACCTGGAGGCC 1628
QY 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
Db 1629 TTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCCTTGGCATTAAACAGCTTC 1688
QY 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
Db 1689 CAAGAAGTGTGGCTTCCGATTAAACAGAACTGGACTGTGATAGCATCTTCTTCTGCCTC 1748
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
Db 1749 CAGCACTGTCCGTATTTGGCGAAATTCGGGTGATGTCAAGGGATCTTCCCAAGAGAT 1808
QY 656 GluSerAlaGluAlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGlu 675
Db 1809 GAGTCCGCTGAGGATGCTCTGTGCTTATGGATGCGGATAAGACCTCATTTAG 1868
QY 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
Db 1869 GAGCAGTGGGAAGATTCTGTCTCATGCTTGGCCACCCACACACCTGCGGCAGCTGGAC 1928
QY 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
Db 1929 CTGGCAGCAGCATCTGACAGAGCGGCCATGAGACCTGTGTGCCAAGCTGAGGCAT 1988
QY 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
Db 1989 CCCACCTGCAAGATACAGACCTGATGTTTAAAGATGCACAGATTACCCCTGTGTGCAG 2048
QY 736 HisLeuTyrArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThr 755
Db 2049 CACCTCTGGAGATCGTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCACC 2108
QY 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
Db 2109 CACCTGAAGGAAGAGATGTAAAGATGGCGTGAAGCCTTAAACACCCCAAAATGTTG 2168
QY 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795
Db 2169 TTGAGTCTTTGAGGCTGGATGTGTGGATGTGCCATGCCATGCTGTACCTGAAGATCTCC 2228
QY 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815
Db 2229 CAATCTCTTACGACTCTCCCGAGCTGAAATCTCTGAGCCTGGCAGGAACAAGGTGACA 2288
QY 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
Db 2289 GACCAGGGAGTAATGCCTCTCAGTGATGCTTGAAGATCTCCAGTGGCCCTGCAGAG 2348
QY 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
Db 2349 CTGATACTGGAGGATGTGGCATCACAGCCACGGGTTGCCAGAGTCTGGCCTCAGCCCTC 2408

QY 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
Db 2409 GTCAGCAACCGGAGCTTGCACACCTGTGCTTATCAACAACAGCCTGGGGAACGAAAGGT 2468
QY 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
Db 2469 GTAAATCTACTGTGTGATCCATGAGGCTTCCCATCTAGTCTGCAGAGGCTGTAGCTG 2528
QY 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915
Db 2529 AATCAGTGCCACCTGGACACGGCTGCTGTGTGTTTCTTGCCTTATGGGTAAAC 2588
QY 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
Db 2589 TCATGGCTGACGACCTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCGTGAAGCTT 2648
QY 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
Db 2649 CTGTGGAGGTCATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGT 2708
QY 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975
Db 2709 CATCTCACCGCCGCTGTGTGAGAGTCTGCTGTGTGATCTCGAGGAGCAGACACCTG 2768
QY 976 LysSerLeuAspLeuThrAspAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
Db 2769 AAGAGCTGTGATCTCAGGACAATGCCCTGGGTGAGCGGTGGGTGTGTCACCTGTGGCAG 2828
QY 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
Db 2829 GGACTGAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGACCTGACT 2888
QY 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
Db 2889 TCTGATTGCTGTGAGGACACTCTCTTGGCCCTTCTCTCAACCCGCACTCTGACAGCTA 2948
QY 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
Db 2949 AACCTGGTGAGAAATTAACCTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCTTTGCC 3008
QY 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
Db 3009 TGTCCCACTGTAACTTACAGATAAATGGCTGTGGAATGGCAGTACCTGTGCAATA 3068
QY 1076 ArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrp 1095
Db 3069 AGGAAGCTGTGGGAAAGTGCAGCTACTCAAGCCCCGAGTCGTAATTCGCGGTAGTTG 3128
QY 1096 HisSerPheAspGluAspArg 1103
Db 3129 CATCTTTTGTGAAGATGACCGG 3152

RESULT 7

US-10-677-943-23
; Sequence 23, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64790
; CURRENT APPLICATION NUMBER: US/10/677,943
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 23

; LENGTH: 3900

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(3600)

US-10-677-943-23

Alignment Scores:

Pred. No.: 0 Length: 3900

Score: 616.00 Matches: 1002

Percent Similarity: 99.21% Conservative: 4

Best Local Similarity: 99.21% Mismatches: 0

Query Match: 45.83% Indels: 8

DB: 12 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-677-943-23 (1-3900)

Qy	98	IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluLeu	117
Db	568	ATTTCAAGCTATGCAACAGAGGTGCCACAGCAGCAGACAGACAGACAGAAATT	627
Qy	118	SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly	137
Db	628	TCACAAGCTATGGAACAGAGGTGCCACAGCAGCAGACAGACAGAAAGGACATGGA	687
Qy	138	GlyAspThrTrpAspTyrIysSerHisValMetThrIysPheAlaGluGluAspVal	157
Db	688	GGTGACACATGGGACTACAGAGTACCGTAGACCAAAATTCGCTGAGGAGGAGATGTA	747
Qy	158	ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla	177
Db	748	CGTCGTAGTTTGAACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	807
Qy	178	PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyIysSerGly	197
Db	808	TTTGATTCAGACCGGTGGGGCTTCGGGCTCGCAGCGTGGTTCGACGGAAGTCAGGA	867
Qy	198	IleGlyIysSerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyLeuTyr	217
Db	868	ATTGGGAATCGGCTCTAGCCAGAGAGTGGTGTGCTGGGGCGCAAGGTGGACTCTAC	927
Qy	218	GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgIysIysGlu	237
Db	928	CAGGGAATGTTCTCTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	987
Qy	238	SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu	257
Db	988	AGCAGTGTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTGCAGCGAG	1047
Qy	258	IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer	277
Db	1048	ATCATGTCCCGACCAAGAGCGTGTGTTCATCATGTGACGGTTTCGATGACCTGGGCTCT	1107
Qy	278	ValLeuAsnAsnAspThrIysLeuCysIysAspTrpAlaGluIysGlnProProPheThr	297
Db	1108	GTCCTCAACATGACACAAAGCTCTGCAAGAGTGGGCTGAGAGCAGCCCTCCGTTACCC	1167
Qy	298	LeuIleArgSerLeuLeuArgIysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1168	CTCATACGCGTCTCTGAGGAAGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1227
Qy	318	ArgAspValGlyThrGluIysLeuIysSerGluValValSerProArgTyrLeuLeuVal	337
Db	1228	AGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTGCTGTCTCTCTCTCTCTCTCTCT	1287
Qy	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis	357
Db	1288	AGAGGAATCTCCGGGGAACAAAGAAATCCACTTGTCTCTCTCTCTCTCTCTCTCTCT	1347
Qy	358	GlnIysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysGly	377

Db	1348	CAGAAACACAAAGGGTTGCGTGCATCAT-CAACAACCGTGAGCTGCTCGACCAAGTCCCA	1406
Qy	377	nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGly	397
Db	1407	GGTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACCTGCGAGCAGTGTGGGGGA	1466
Qy	397	uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG	417
Db	1467	GAGCGTCGCCCTTCAACCAACAGCTCACAGGCTCACAGGCTCGACGCGCTTTTGC-GTTTCATC	1525
Qy	417	lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL	437
Db	1526	AGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTCGAGGAAGAGTGTCTCTCGA	1585
Qy	437	ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgIysSerValPheAspGlyA	457
Db	1586	AGCGTCTCTCCGCTATGGCTGTGAGGAGTGTGAATAGAAAGTCAAGTGTGTGTGTGTGT	1645
Qy	457	spAspLeuMetValGlnGlyLeuGlySerGluLeuArgAlaLeuPheHisMetAsnI	477
Db	1646	ACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGCTGCTCTGTTTCACATGAACA	1705
Qy	477	leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA	497
Db	1706	TCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTCCAGG	1765
Qy	497	spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP	517
Db	1766	ACTTCTGTGCGGCTGTACTAGTGTGAGGCGCTGGAAATCGAGCCAGCTCTCTGCCC	1825
Qy	517	roLeuTyrValGluIysThrIysArgSerMetGluLeuIysGlnAlaGlyPheHisIleH	537
Db	1826	CTCTGTACGTTGAGACACAAAGAGTCCATGAGCTTTAAACAGCAGGCTTCCATATCC	1885
Qy	537	isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL	557
Db	1886	ACTCGCTTTGGATGAGCGTTTCTTGTGTCCTCGTGAGCGAAGAGCTTAAGGAGGCGAC	1945
Qy	557	euGluValLeuLeuGlyCysProValProLeuGlyValIysGlnLeuLeuHisTrpV	577
Db	1946	TGGAGGTCCTGTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGG	2005
Qy	577	alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH	597
Db	2006	TCTCTCTGTTGGGTCAGCAGCCTTAATGCCACACCCCGAGGAGACACCTCGAGCGCTTC	2065
Qy	597	isCysLeuPheGluThrGlnAspIysGluPheValArgLeuAlaLeuAsnSerPheGlnG	617
Db	2066	ACTGCTCTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACAGAGCTTCCAAG	2125
Qy	617	luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH	637
Db	2126	AAGTGTGGCTTCGATTAACAGAACCTGGACTTGATAGATCTTCTCTCTGCTCCAGC	2185
Qy	637	isCysProTyrLeuArgIysIleArgValAspValIysGlyIlePheProArgAspGluS	657
Db	2186	ACTGTCGCTATTTCGGGAAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGATGAGT	2245
Qy	657	erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG	677
Db	2246	CCGCTGAGGCGATGTCTGTGCTGCTCTATAGGATCGGGATAGACCTCATTTGAGGAGC	2305
Qy	677	lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG	697
Db	2306	AGTGGGAAGATTTCTGCTCCATGCTTGGCACCCACACACCTCGCGCAGCTGACCTGG	2365
Qy	697	lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT	717
Db	2366	GCAGCAGATCTTCAGACAGAGGGGCCATGAAGACCTGTGTGCCAGCTGAGGATCCCA	2425
Qy	717	hrCysIysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL	737

Db 2426 CCTGCAAGATACAGACCCTGATGTTTACAAATGACACGATTAACCCCTGGTGTGCAACACC 2485
Qy 737 euTrpArgIleValMetAlaIenArgAenLeuArgSerLeuAsnLeuGlyThrHisL 757
Db 2486 TCTGGAGAATGTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCCACC 2545
Qy 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAGAGATGTAAGGATGGCGTGAAGCTTTAAACACCCCAAAATGTTTGTGG 2605
Qy 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCCTTCAGGCTGGATGCTGTGATGACCCATGCGCTGTACTGAAGATCTCCCAA 2665
Qy 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCCTTACGACCTCCCCACGCTGAATCTCTGAGCCTGGCAGGAACAAGGTGACAGACC 2725
Qy 817 InGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGAGATTAAC-GCCTCTCAGTATGCTTGAAGGTCTCCCAAGTGGCCCTGCAGAAGCTG 2784
Qy 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGGACTGTGGCATCACAGCCACGAGTTGCCAGAGTCTGGCCTCAGCCCTGTC 2844
Qy 857 SerAenArgSerLeuThrHisLeuCysLeuSerAenAenSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACCGAGCTTGACACACTGTGCTTATCCAAACACAGCTGGGGACCAAGAGTGTA 2904
Qy 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTGCATCATGAGGCTTCCCCACCTGTAGTCTGCAGAGGCTGATGCTGAAT 2964
Qy 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGGCCACCTGGACACGGCTGGCTGGG-TTCTCTTGACCTTGGCTTATGGGTAACTC 3023
Qy 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAenGlyValLysLeuLe 936
Db 3024 ATGCTGTGACGACCTTGAGCCTTAGCATGAACCTGTGGAAGACAAATGGCGTGAAGCTTCT 3083
Qy 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGGAGGTCTATGAGAGAACCATTTGTTCATCTCCAGGACTGGAGTTGGTAAGTGTCAT 3143
Qy 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
Db 3144 TCTCACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAA 3203
Qy 976 sSerLeuAspLeuThrAspAenAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGl 996
Db 3204 GAGCCTGTGATCTCAGGACAAATGCCCTGGGTGAGCGTGGGTGCTGCTGCTGCGAGGG 3263
Qy 996 yLeuLysGlnLysAenSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe 1016
Db 3264 ACTGAGCAAAAGAACAGTGTCTGACAGACTCGGGTTGAGGCATGTGACACTTC 3323
Qy 1016 rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs 1036
Db 3324 TGATTGCTGTGAGGACACTCTCTTGGCCCTTTCTGTCAACCGGCATCTGACCACTTAA 3383
Qy 1036 nLeuValGlnAenAenPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy 1056
Db 3384 CCTGGTGAGAAATAACTTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCGCTG 3443
Qy 1056 sProThrSerAenLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleAr 1076
Db 3444 TCCACAGTCTAACTTACAGATAATGGCTGTGGAATGGCAGTACCTGTGCAAAATAG 3503
Qy 1076 gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi 1096
Db 3504 GAAGCTGTGGAGGAAGTGACAGCTACTCAAGCCCGAGTCTGTAATTGACCGTAGTTCGCA 3563

Qy 1096 sSerPheAspGluAspAspArg 1103
Db 3564 TTCTTTTGTGATGAAGATGACCGG 3585

RESULT 8

US-10-399-443-23
; Sequence 23, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; Secretary, Department of Health & Human Services, The National Institute of
; Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 3900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3600)
; OTHER INFORMATION:
US-10-399-443-23

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-399-443-23 (1-3900)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCAACAGCTATGACACAGAGAGTGGCCACAGCAGCAGACAGACAGAGAAAT 627
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 628 TCACAAAGCTATGAAACAAGAGAGTGGCCACAGCAGCAGACAGACAGAAACAAGGACATGGA 687
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTACAGAGTCAAGAGTCACTGCTGCTGACCAAAATTCGCTGAGGAGGAGTGA 747
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTGTAGTATTTGAAAACACTGCTGCTGACTGCTGCCGGAATGCAACGTTGGCTGGTCT 807
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 808 TTTGATTACAGCCGCTGGGCTTCCGCCCTCGCACGCTGGTTCGACCGAAAGTCAGGA 867
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyr 217
Db 868 ATTTGGAAATCGGCTCTAGCCAGAGAGTCTGCTGCTGCTGGCGCGCAAGGTGACTCTAC 927
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237

Db 928 CAGGGAATGTTCTCTACGTCTTCTTCCTCCCGTTAGAGAGATCAGCGGAAGGAG 987
Qy 238 SerSerValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGlu 257
Db 988 AGCAGTGTACAGAGATTCTCTCAGGAGGTGGCCAGACTCCAGGCTCCGGTACGGAG 1047
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 1048 ATCATGTCCGACCAAGAGGCTGTGTTCATCATTTGACGGTTTCGATGACTGGGCTCT 1107
Qy 278 ValLeuAsnAsnAspThrIleLeuCysIleAspTyrAlaGluLeuGlnProPheThr 297
Db 1108 GTCCCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAGAACAGCCCTCCGTTACC 1167
Qy 298 LeuIleArgSerLeuLeuArgIleValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACGAGTCTGCTGAGGAGGTCTCTGCTCCCTGAGTCTTCTGATCGTCAACGTC 1227
Qy 318 ArgAspValGlyThrGluIleLeuIleValSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTCTGTCTCCCGTTACTCTGTAGTT 1287
Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisIleLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTTGAAGCGGGATTTGGTGAAGAT 1347
Qy 358 GlnIleThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1348 CAGAGACACAGGCTTGGTGGATCAT-CAACAAACCGTGAGCTGCTCGAACAGTGCCA 1406
Qy 377 nValProIleValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGln 397
Db 1407 GGTGCGCCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGA 1466
Qy 397 uSerValIleProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db 1467 GAGCGTCGCCCCCTTCACACCAACCGCTCACAGGCTGCAGCCCGCTTTTTC- GTTTTCATC 1525
Qy 417 InLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db 1526 AGCTCACCCCTCGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTGA 1585
Qy 437 ysArgPheCysArgMetAlaValGluGlyValTyrProAsnArgIleSerValPheAspGly 457
Db 1586 AGCGCTTCTCCCGTATGCTGTGGAGGAGGTGGAAATAGGAAGTCACTGTTTGTGATGTG 1645
Qy 457 spAspLeuMetValGlnGlyLeuGlyGlyLeuSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db 1646 ACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTCTGTGTTTCACATGAACA 1705
Qy 477 leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA 497
Db 1706 TCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGG 1765
Qy 497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db 1766 ACTTCTGTGGCCCTTGTACTACGTGTTAGAGGGCTCGAAATCGAGCAGCTCTCTGCC 1825
Qy 517 roLeuTyrValGluIleThrIleArgSerMetGluLeuIleGlnAlaGlyPheHisIleH 537
Db 1826 CTCGTAGCTGTGAGAAGCAAGAGGTCCATGGAGCTTAAACAGGCGAGGCTTCCATATCC 1885
Qy 537 isSerLeuTyrMetIleValArgPheLeuGlyLeuValSerGluAspValArgArgProL 557
Db 1886 ACTCCGCTTGGATGAGCGTTCTGTTTGGCTCTGAGCGAAGACGTAAAGAGGCCAC 1945
Qy 557 euGluValLeuLeuGlyCysProValProLeuGlyValIleGlnIleLeuHisIleH 577
Db 1946 TGGAGGCTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAACAGAGCTTCTGCACTGGG 2005
Qy 577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db 2006 TCTCTCTGTGGGTGAGAGCCCTAATGCGCACCCAGGAGACACCCCTGGACGCTTCC 2065

Qy 597 isCysLeuPheGluThrGlnAspIleValArgLeuAlaLeuAsnSerPheGlnG 617
Db 2066 ACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGACTTCCAA 2125
Qy 617 luValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db 2126 AAGTGTGGCTTCCGATTAAACAGAACTGGACTTGTATAGCATCTTCTTCTGCTCCAGC 2185
Qy 637 isCysProTyrLeuArgIleArgValAspValIleGlyIlePheProArgAspGluS 657
Db 2186 ACTGTCCGTATTTCCGAAAATTCCGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGT 2245
Qy 657 erAlaGluAlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGluGluG 677
Db 2246 CCGCTGAGGACATGTCTGTGGTCCCTCTATGTGATCGGGATAGACCTCATTCAGGAGC 2305
Qy 677 InTyrGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db 2306 AGTGGGAAGATTTCTGCTCCATGTCTGGCACCCACCCACACCTCGCGCAGCTGGACCTGG 2365
Qy 697 lySerSerIleLeuThrGluArgAlaMetIleThrLeuCysAlaLeuLeuArgHisProT 717
Db 2366 GCAGCAGATCTCTACAGAGCGGCCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCCA 2425
Qy 717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db 2426 CCTGCAGATACAGACCTCTGATGTTTGTAGAAATGCACAGATATACCCCTGGTGTCAACACC 2485
Qy 737 euTyrArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db 2486 TCTGGAAATCGTCTATGCCCAACCGTAACCTAAGATCCCTCAACTTGGGAGCACCCACCC 2545
Qy 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProIleCysLeuLeuG 777
Db 2546 TGAAGGAAGAGGATGTAAAGATGGCGTGTGAAGCCTTTAAACACCCCAAAATGTTTGTGG 2605
Qy 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCTTTGAGGCTGGATTTGCTGTGGATTGACCCATGCTGTACTTACCTGAGAGTCTCCCAA 2665
Qy 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCTTACGACCTCCCCAGGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACC 2725
Qy 817 InGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGAGTAAC-GCCTCTCAGTGATGCCCTTGAGGGTCTCCAGTGCGCCCTGCAGAAGCTG 2784
Qy 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGGCTCAGCCCTCGTC 2844
Qy 857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACCGGAGCTTGACACACCTGTGCTATCCAAACAACAGCCTGGGGAACAAGAGTGTA 2904
Qy 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTGATCATCATGAGGCTTCCCCACTGTGTAGTGTGACAGAGGCTGATGCTGA 2964
Qy 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGCCACCTGGACACGGCTGGCTGTGG-TTCTCTTGCATCTGGCTTATGGGTAACTC 3023
Qy 916 rTyrLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe 936
Db 3024 ATGGCTGACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAAATGGCGTGAAGCTTCT 3083
Qy 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGCGAGGTCTATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTC 3143


```
OY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 10
US-10-399-443-1
; Sequence 1, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-1

Alignment Scores:
Pred. No.: 0 Length: 1157
Score: 385.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.65% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x US-10-399-443-1 (1-1157)

OY 252 GlnAlaProValThrGluLeuMetSerArgProGluArgLeuLeuPheIleAspGly 271
Db 1 CAAGCTCCGGTGACGGAGATCATGTCGCCGACAGAAAGCGTGTGTTTCATCATTCGCGT 60
OY 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrrAlaGlu 291
Db 61 TTCGATGACCTGGGCTCTGTCTCAACAATGACACAAAGCTCTGCACAAAGACTGGGCTGAG 120
OY 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
Db 121 AAGCAGCTTCGGTTCACCTCATAGCAGTCTGCTGAGGAAGGTCTGCTCTCCCTGAGTCC 180
OY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCAAGAGAGCTCAAGTCAGAGGTCGTGTCT 240
OY 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351
Db 241 CCCCGTTACCTGTTAGTAGAGGAATCTCCGGGGAACAAGAATCCACTTCCTCTTTGAG 300
OY 352 ArgGlyIleGlyHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
Db 301 CCGGGGATTTGTGACATCAGACAGACAGAGGTTCGTGCGATCATGACAAACCGTGAG 360
OY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACAGTGCAGCGTCCCGCCCTCTCTCATCTGCGTGGCCCTGAGCTG 420
OY 392 GlnAspValValGlyLysSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
Db 421 CAGGACGTGGTGGGGAGAGCGTCCGCCCTTCAACCAAGAGCTCACAGGCTGACGCC 480
```

```
OY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
Db 481 GCTTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCATCTGGAG 540
OY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrrAsnArgLys 451
Db 541 GAAAGAGTTGTCTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGAATAGGAAG 600
OY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
Db 601 TCAGTGTGTTGACGGTGACGACCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCT 660
OY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTrrThrPhePhe 491
Db 661 CTGTTTCACATGAACATCTCTTCCCAGACAGCCACTGTGAGGAGTACTACACCTTTCTC 720
OY 492 HisLeuSerLeuGlnAspPheCysAlaLeuTyrTrrValLeuGluGlyLeuGluIle 511
Db 721 CACCTCAGTCTCCAGGACTTCTGTGCCGCTTGTACTACGTGTGTAGAGGCGCTGGAATC 780
OY 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTAGCTTGAGAAAGACAAAGAGGTCCATGGAGCTTAAACAG 840
OY 532 AlaGlyPheHisIleHisSerLeuTrrMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCCACTCGCTTGGATGAAGCGTTCCTTGTGGCTCTGAGGCGAA 900
OY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GACGTAAGGAGGCGCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCTCCCTGGGGTGAAGCAG 960
OY 572 LysLeuLeuHisTrrValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGACCTGGGTCTCTCTGTTGGGTGACAGCGCTTAATGCCACCCAGGAGAC 1020
OY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGACGCTTCCACTGTCTTTTCAGACTCTCAGACAAAGAGTTGTTCGCTTGGCA 1080
OY 612 LeuAsnSerPheGlnGluValTrrProIleAsnGlnAsnLeuLeuAspLeuIleAlaSer 631
Db 1081 TTAACAGCTTCAAGAAAGTGTGCTTCCGATTAAACCAGAACCTGGACTTGTATGATCATCT 1140
OY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 11
US-10-677-943-3
; Sequence 3, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64790
; CURRENT APPLICATION NUMBER: US/10/677,943
; CURRENT FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1075
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-677-943-3

Alignment Scores:
Pred. No.:      8,37e-192      Length:      1075
Score:          196.00         Matches:      252
Percent Similarity: 99.21%     Conservative:  0
Best Local Similarity: 99.21%  Mismatches:      1
Query Match:      14.58%      Indels:        2
DB:              12          Gaps:          0

US-10-066-521-6 (1-1344) x US-10-677-943-3 (1-1075)

QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGCCTATCCCAACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAGAGGTGTAATCTACTGTGTCGATCCATGAGGCTTCCCACTGTAGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCGCTGGCTGGCTGGT-TTCTCTTGGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGGCGTTATGGGTAACTCATGTGCTGACGCACTTAGCATGAACCCCTGTGGAAGA 240
QY 930 pAsnGlyVallyLeuLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACCACTTGTCTCATCTCCAGGACCT 300
QY 950 uGluLeuVallyCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTGTGTAAATGTCTATCTACCGCGCGTGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGGAAGACCTTGGATCTCAGGCAATGCGCCCTCCCAAGGAATGATGAAGCT 420
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCAATGTGGACTGACTTCTGATTCGTGTGAGGCACTCTCTTGGCCCTTTCTTGCACACCG 540
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACCAGCTAAACCTGGTGCAGATAAATTCAGTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
Db 601 GTGTTCGGCCCTTTCCTGTGTCACAGTCTAACTTACAGATAATTGGGTGTGGAATGGCA 660
QY 1070 nTyrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLysProArgValVa 1090
Db 661 GTACCCCTGTGCAANTAGGAGCTGCTGGAGGAAGTGCAGTACTCAAGCCCGGAGTCTGT 720
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTATGTTGGCAATCTTTTGTATGAGATGACCGG 760
```

RESULT 12

```
US-10-399-443-3
; Sequence 3, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; SECRETARY: Secretary, Department of Health & Human Services, The National Institute
```

```
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-3

Alignment Scores:
Pred. No.:      8,37e-192      Length:      1075
Score:          196.00         Matches:      252
Percent Similarity: 99.21%     Conservative:  0
Best Local Similarity: 99.21%  Mismatches:      1
Query Match:      14.58%      Indels:        2
DB:              13          Gaps:          0

US-10-066-521-6 (1-1344) x US-10-399-443-3 (1-1075)

QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGCCTATCCCAACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAGAGGTGTAATCTACTGTGTCGATCCATGAGGCTTCCCACTGTAGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCGCTGGACACGGCTGGCTGGT-TTCTCTTGGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGGCGTTATGGGTAACTCATGTGCTGACGCACTTAGCATGAACCCCTGTGGAAGA 240
QY 930 pAsnGlyVallyLeuLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACCACTTGTCTCATCTCCAGGACCT 300
QY 950 uGluLeuVallyCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTGTGTAAATGTCTATCTACCGCGCGTGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGGAAGACCTTGGATCTCAGGCAATGCGCCCTCCCAAGGAATGATGAAGCT 420
QY 990 lAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLys 1010
Db 421 TGTCTGCGCTGTGGAGGAGCTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAA 480
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCAATGTGGACTGACTTCTGATTCGTGTGAGGCACTCTCTTGGCCCTTTCTTGCACACCG 540
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACCAGCTAAACCTGGTGCAGATAAATTCAGTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
Db 601 GTGTTCGGCCCTTTCCTGTGTCACAGTCTAACTTACAGATAATTGGGTGTGGAATGGCA 660
QY 1070 nTyrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLysProArgValVa 1090
Db 661 GTACCCCTGTGCAANTAGGAGCTGCTGGAGGAAGTGCAGTACTCAAGCCCGGAGTCTGT 720
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTATGTTGGCAATCTTTTGTATGAGATGACCGG 760
```

```
Db      601 GTGTTGGCCTTGCTGTCCACGCTTAACCTTACAGATAAATGGGCTGTGGAATGGCA 660
Qy      1070 nTyrProValGlnIleAArgLysLeuGluGluValGlnLeuLysProAArgValVa 1090
Db      661 GTACCTGTGCAAAATAAGGAAGCTCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGT 720
Qy      1090 lIleAspGlySerTrpHisSerPheAspGluAspAspArg 1103
Db      721 AATTGACGGTAGTTGGCATTCTTTTGTATGAAGATGACCGG 760

RESULT 13
US-10-027-632-258159
; Sequence 258159, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258159
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258159

Alignment Scores:
Pred. No.:      4,92e-48      Length:      2099
Score:          57.00      Matches:      57
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    4.24%      Indels:      0
DB:             16      Gaps:      0

US-10-066-521-6 (1-1344) x US-10-027-632-258159 (1-2099)
Qy      668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db      220 ATGCGGATAGACCCCTCATTGAGGAGCAGTGGGAGAGATTCTGCTCCATGCTTGGCACC 279
Qy      688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetIys 707
Db      280 CACCCACACCTGCGGAGCTGGACCTGGGAGCAGCATCTGCAGAGCGGGCCATGAAG 339
Qy      708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet 724
Db      340 ACCCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATACAGACCCCTGATG 390

RESULT 15
US-10-029-386-10957/c
; Sequence 10957, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GEN
; EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10957
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008753.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.37
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: B1911853.1, EVALUAE 5.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P33076, EVALUAE 1.00e-09
; OTHER INFORMATION: NT HIT: g114731965, EVALUAE 1.90e+00
```

```
Db      601 GTGTTGGCCTTGCTGTCCACGCTTAACCTTACAGATAAATGGGCTGTGGAATGGCA 660
Qy      1070 nTyrProValGlnIleAArgLysLeuGluGluValGlnLeuLysProAArgValVa 1090
Db      661 GTACCTGTGCAAAATAAGGAAGCTCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGT 720
Qy      1090 lIleAspGlySerTrpHisSerPheAspGluAspAspArg 1103
Db      721 AATTGACGGTAGTTGGCATTCTTTTGTATGAAGATGACCGG 760

RESULT 13
US-10-027-632-258159
; Sequence 258159, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258159
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258159

Alignment Scores:
Pred. No.:      4,92e-48      Length:      2099
Score:          57.00      Matches:      57
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    4.24%      Indels:      0
DB:             16      Gaps:      0

US-10-066-521-6 (1-1344) x US-10-027-632-258159 (1-2099)
Qy      668 MetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThr 687
Db      220 ATGCGGATAGACCCCTCATTGAGGAGCAGTGGGAGAGATTCTGCTCCATGCTTGGCACC 279
Qy      688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetIys 707
Db      280 CACCCACACCTGCGGAGCTGGACCTGGGAGCAGCATCTGCAGAGCGGGCCATGAAG 339
Qy      708 ThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMet 724
Db      340 ACCCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATACAGACCCCTGATG 390

RESULT 14
US-10-027-632-258159
; Sequence 258159, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
```

US-10-029-386-10957

Alignment Scores:

Pred. No.:	0.0365	Length:	509
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	15	Gaps:	0

US-10-066-521-6 (1-1344) x US-10-029-386-10957 (1-509)

Qy 262 ProGluArgLeuPheIleAspGlyPheAsp 273

Db 178 CCCGAGCGCTCTTTTCATCATCGCGCTTCGAT 143

Search completed: July 16, 2004, 18:42:58

Job time : 1422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 15, 2004, 08:50:58 ; Search time 12090 Seconds
(without alignments)
4818.283 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 7074
Sequence: 1 MEGDKSLTFSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10066521/runat_13072004_122212_11081/app_query.fasta_1.1543
-DB=genEmbl -QWMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn 1 1 8359 @runat_13072004_122212_11081 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sv:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sv:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5763.5	81.5	3885	9	AY154460 Homo sapi
2	5756	81.4	3489	6	AX478549 Sequence
3	5756	81.4	3926	6	AX704821 Sequence
4	5741.5	81.2	3885	9	AY054986 Homo sapi
5	5741.5	81.2	3900	6	AX427610 Sequence
6	5724.5	80.9	5859	6	AX459881 Sequence
7	5724.5	80.9	6939	6	AX459873 Sequence
8	5724.5	80.9	6939	6	AX459891 Sequence
9	5683.5	80.3	3830	6	AX704823 Sequence
10	2814	39.8	157141	9	AC011470 Homo sapi
11	2814	39.8	193609	9	AC024580 Homo sapi
12	2695	38.1	3470	10	AY196362 Mus muscu
13	2692	38.1	3405	10	AY329487 Mus muscu
14	2688	38.0	3447	6	AX427592 Sequence
15	2688	38.0	3447	10	AF074018 Mus muscu
16	2688	38.0	3453	10	AY329486 Mus muscu
17	2688	38.0	3480	10	AY196361 Mus muscu
18	2688	38.0	3534	10	BC053384 Mus muscu
19	2686	38.0	3454	10	AY329490 Mus muscu
20	2685	38.0	3432	10	AY329484 Mus muscu
21	2567	36.3	3375	10	AY329489 Mus muscu
22	2553.5	36.1	3372	10	AY329485 Mus muscu
23	2540	35.9	3303	10	AY329488 Mus muscu
24	2533	35.8	3303	10	AY329491 Mus muscu
25	2005	28.3	1157	6	AX427588 Sequence
26	1885.5	26.7	3282	9	AY154469 Homo sapi
27	1885	26.6	193662	2	AC127352 Mus muscu
28	1732	24.5	167509	2	AC012107 Homo sapi
29	1435	20.3	3333	9	AF442488 Homo sapi
30	1435	20.3	3339	9	AY072792 Homo sapi
31	1435	20.3	3368	9	AF479747 Homo sapi
32	1434	20.3	3365	9	BC050326 Homo sapi
33	1396	19.7	3234	9	AF482706 Homo sapi
34	1356	19.2	231371	2	AC132633 Rattus no
35	1356	19.2	323826	2	AC110429 Rattus no
36	1335	18.9	1075	6	AX427590 Sequence
37	1331.5	18.8	3132	9	AY154468 Homo sapi
38	1331.5	18.8	3300	6	AX459869 Sequence
39	1320.5	18.7	3108	6	AX417214 Sequence
40	1318	18.6	1576	10	F143559S07
41	1316.5	18.6	3190	6	AX714395 Sequence
42	1316.5	18.6	3190	9	AK056688 Homo sapi
43	1315.5	18.6	3172	6	AX743691 Sequence
44	1315	18.6	3731	9	AY116204 Homo sapi
45	1306.5	18.5	3507	9	BC028069 Homo sapi

ALIGNMENTS

RESULT 1

AY154460 3885 bp mRNA linear PRI 20-FEB-2003
LOCUS
DEFINITION Homo sapiens NALP5 (NALP5) mRNA, complete cds.
ACCESSION AY154460
VERSION AY154460.1 GI:28436363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3885)
AUTHORS Tschopp, J., Martinon, F. and Burns, K.
TITLE NALP5: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
MEDLINE 2451042
PUBMED 12563287
REFERENCE 2 (bases 1 to 3885)
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
source
1..3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3885
/gene="NALP5"
1..3603
/gene="NALP5"
/notes="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains, similar
to mouse WATER protein"
/codon_start=1
/product="NALP5"
/protein_id="AA018156.1"
/db_xref="GI:28436364"
translation="MKVAGLELGALLSAPALVTLSTGPTCSILPKNPLPPQNL
SSQCIKMGDKSITFSYGLQWCLYELDKREFQTFKELUKKSSSTTCSIFQFELE
NANVECLALLHEYGASLAWAYSISIFENNNLTSEKARDDMKRHSPPEDPATMTD
QGPSKVPQISQVQDSATAETKEQEI SQAMEQEGATAETEEQEI SQAMEQEG
TAATEBQSGGDTWYKSHVMTKFAEEEDVRRSFENATAADPEMOTLAGAFSDRWG
FRPTVYHKGSGIKGALARRIVLCWAQGLYQGMFSYVFLPVRMORKKSSVTE
FLSEWPDQAPVTEINSRPERLLFIIDGFDDLGSLVNDTLCKDKWAQKQPPETLIR
SLRKVLPSFLIVTRDVGTEKLKSEVSPRYLLVRLGSGEQRHLLERIGIEHQ
KTQGLRAMNRELDDQCPVAVGSLICVALQDQVVGVSAPFNQTLGLHAAPFVPH
QLTPRGVVRRLNLEERVLKRCFMAVEGVNMRKSVDFGDDMLVQGLSELRALFH
MNIILPDSHCEEYVTFPHLSLQDFCAALYVLELEI EPALCPYVETKRSMLKQA
GFHITSLWMKRFGLVSEVDRRLPVLVLLGCPVLGVKQKLLHWVSLGOQPNATPG
DTLDAFLCPETQDKEFVRLALNSFQSWLPINQNLDLIASSFCLOHCPVLRKRDV
KGIFPDESABEACPVPLMRDKTLIEQWEDFCMLGTHPHLRQLDLGSSILITERAM
KTLCAKLRHPTCKITLMFRNAQITPGVHLRIVMANRLNSLNLGGTHLKEEDVRM
ACEALHPKCLLESRLDCCGLTHACYLKI SOLITSPSLKSLSLAGNKVTDQGVNPL
SDALRVQCALQKLLIEDCGITATGCOSLASALVSNRSLTHLCLSNNSLGNQGVNLLC
RSMELPHCSLQRLMLNCHLDTAGCGFLALMGNSWLTHLSLSPNPVEDNGVKLLCE
VMRPSCHLDLQVVKHLTAACCESLSVLSRSLHLSLDLTDNALDGDGVAAALCEG
LKQNSVTLRLGLKACGLTSDCCESLSALUSCNRLSLNLVQNNLSPFGMKMLKCSAF
ACPTSNLQIITGLMKWQPVQIRKLLLEEVQLLKPVRVLDGSMHSDFEDDDRYWKN"

Alignment Scores:
Pred. No.: 0 Length: 3885
Score: 5763.50 Matches: 1102
Percent Similarity: 96.42% Conservative: 2
Best Local Similarity: 96.24% Mismatches: 0
Query Match: 81.47% Indels: 41
DB: 9 Gaps: 1

US-10-066-521-6 (1-1344) x AY154460 (1-3885)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTyrCysLeuTyrGlu 20
Db 154 ATGGAAGAGAGACAAATCGCTTCCACCTTTTCCAGCTACCGGCTGCAATGGTGTCTCTATGAG 213

Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 214 CTAGACAAGAGAAATTTTCAGACATTTCAAGGAATTTACTAAAGAGAAATCTTCAGAAATCG 273

Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 274 ACCACATGCTCTATTTCACAGTTTGAATCGAAGATGCCAACGTGGNATGCTTGGCACTC 333

Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrAlaThrSerIleSerIlePheGlu 80
Db 334 CTCTTGTCATGAGTATTATGAGCATCGTGGCTGGCTACGTCACATTAGCATCTTTTGA 393

Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys----- 97
Db 394 AACATGAACCTGCGAACCTCTCGGAGAGGCGCGGATGACATGAAAGACATTCACCA 453

Qy 97 ----- 97

Db 454 GAAGATCCTGAAGCAAGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCAGGAATT 513

Qy 98 -----IleSer 99

Db 514 TCACAAAGCTGTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAACAAATTTCA 573

Qy 100 GluAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGluSerGln 119
Db 574 CAAGCTATGGAAACAAGAGGTGCCACAGCAGCAGACAGACAGAAAGAAATTTCA 633

Qy 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAsp 139
Db 634 GCTATGGAAACAAGAGGTGCCACAGCAGCAGACAGACAGACAGACATGGAGGTGAC 693

Qy 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg 159
Db 694 ACATGGAGCTTACAAGAGTCACGTGATGACCAATTCGCTGAGGAGGAGGATGTACGTCGT 753

Qy 160 SerPheGluAsnThrAlaAlaAspTyrProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db 754 AGTTTTGAAACACTGCTGCTGACTGGCGGAAATGCAACGTTGGCTGGCTTTGAT 813

Qy 180 SerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCCGGCTTCGACCGTGGTCTGCACCGGAAAGTCAGAAATGGG 873

Qy 200 LysSerAlaLeuAlaArgArgIleValLeuCysTyrTrpAlaGlnGlyGlyLeuTyrGlnGly 219
Db 874 AAATCGGCTTAGCCAGAGGATCGTGTGCTGGCGCGCAAGGTGGACTCTACCAAGGA 933

Qy 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACCTTCTTCTCTCCCGTTAGAGATGCGCGGAGAAAGAGAGAGAGCAGT 993

Qy 240 ValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTACAGAGTTCACTCCAGGAGTGCACAGACTCCAGGCTCCGTCGTCGCGAGATCATG 1053

Qy 260 SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAGGCTGTTGTTTCATTCATTCAGCGTTTCGATGACCTGGGCTCTGCTC 1113

Qy 280 AsnAsnAspThrLysLysLysCysLysAspTyrAlaGluLysGlnProPheThrLeuLeu 299
Db 1114 AACAAATGACAAAGAGCTCTGCAAGAGACTGGCTGGAGAGCAGCCTCGTTTACCCCTCATA 1173

Qy 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGTCTGCTGAGGAGGTCCTGCTCCCTGAGTCTCTCTGATCGTCCCGTCAGAGAC 1233

Qy 320 ValGlyThrGluLysLysLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCACAGAGAGAGCTCAAGTCAGAGTGTGCTCTCCCGTGTACCTGTTAGTTAGAGGA 1293

QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 1294 ATCTCGGGGAAACAAGAATCCACTTGTCTTGGCGCGGATTGGTGAGCATCAGAAG 1353
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlnLeuLeuLeuAspGlnCysGlnValPro 379
DB 1354 ACACAGGGTTGCTGGCCATCATGAACAACCGTAGCTGCTCGACCCAGTGCAGGTGCC 1413
QY 380 AlaValGlySerLeuIleCysValAlaLeuLeuGlnLeuGlnAspValValGlyGluSerVal 399
DB 1414 GCCGTGGGCTCTCTCATCTGCTGCGCTGCAGCTGCAGGACGTGGTGGGGAGAGCGTC 1473
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 1474 GCCCCCTTCAACCAACACTCACAGGCTGCACGCGCTTTGTGTGTTTCATCAGCTCACC 1533
QY 420 ProArgGlyValValArgCysLeuAsnLeuGluArgValValLeuLysArgPhe 439
DB 1534 CCTCGAGCGGTGGTCCGCGCTCTCAATCTGGAGGAAGAGTTGCTTGAAGCGCTTC 1593
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu 459
DB 1594 TGGCGTATGGCTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCAGCGGTGACGACCTC 1653
QY 460 MetValGlnGlyLeuGlyGluSerGlnLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 1654 ATGGTTCAAGGACTCGGGGAGTCTGAGCTCGCTGCTCTGTGTTTCATCATGAACATCTCTC 1713
QY 480 ProAspSerHisCysGluGluTrpThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 1714 CGAGACGCACTGTGAGGAGTACTACACCTTCTCCACCTCAGTCTCCAGGACTTCTGT 1773
QY 500 AlaAlaLeuTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 1774 GCCGCTTGTAAGTGTAGAGGCTTGAATCGAGGCGAGCTCTTGCCTCTGTAC 1833
QY 520 ValGlnLysThrLysArgSerMetGlnLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 1834 GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCACTCGCTT 1893
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 1894 TGGATGAAGCGTTCTGTTTGGCTCTGTGAGGACGACGTAAGAGGCGCCACTGGAGGTC 1953
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 1954 CTGCTGGGCTGCCGTTCCTCGGGGTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTG 2013
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 2014 TTGGGTCAAGCAGCCTAATGCCACCAACCCAGGAGACACCTTGGAGCGCTTCCACTGTCTT 2073
QY 600 PheGlnThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 2074 TTCGAGACTCAAGACAAGAGTTGTTCGCTTGGCATTTAAACAGCTTCCAAGAGAGTGTGG 2133
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 2134 CTTCGGATTAACGAGACCTGGACTTGTATAGCATCTTCTTCTGCTCCAGCACTGTCCG 2193
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 2194 TATTTGGGAAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCCGCTGAG 2253
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrIleGluGlnTrpGlu 679
DB 2254 GCATGCTCTGTGGTCCCTCTATGATGCGGATAGACCCCTCATTTGAGGAGCAGTGGGAA 2313
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 2314 GATTTCTGCTCATGCTGGGACCCACCACTTGGCGAGCTGGACCTGGGCGAGCAGC 2373
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719

DB 2374 ATCTCGACAGAGGGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCCTGCAAG 2433
QY 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
DB 2434 ATACAGACCTTGATGTTTAGAAATGCACATTAACCTCGTGGTGGCAGCACCTCTGAGA 2493
QY 740 IleValMetAlaAsnArgAsnLeuLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
DB 2494 ATCGTCATGCCAACCGTAACCTAAGATCCTCAACTTTGGAGGCACCCACCTGAAGGAA 2553
QY 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
DB 2554 GAGGATGAAGGATGGCGTGTGAAGCCTTAAACACACCCCAAATGTTTGTGGAGTCTTTG 2613
QY 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTrpLeuLysIleSerGlnIleLeuThr 799
DB 2614 AGGCTGATTTGCTGTGATTTGACCCATGCTGTACCTGAAGATCTCCCAAAATCTTTACG 2673
QY 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
DB 2674 ACCTCCCCAGCCTGAAATCTCTGAGCCTGCGAGGAAACAAGGTGACAGACCCAGGAGTA 2733
QY 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
DB 2734 ATGCCTCTCAGTGATGCTTTGAGAGTCTCCAGTGCGCCCTGCAGAAAGCTGATCTGGAG 2793
QY 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
DB 2794 GACTGTGGCATACAGCACGGTTGCCAGTCTGSCCTTCAGCCCTCTGCAGCAACCGG 2853
QY 860 SerLeuThrHisLeuCysLysLeuSerAsnAsnSerLeuGlyAsnGlyValAsnLeuLeu 879
DB 2854 AGCTTGACACACCTGTGCCTATCCAAACAACAGCCTGGGGAACGAAGGTGTAATCTACTG 2913
QY 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
DB 2914 TGTGATTCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCAC 2973
QY 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
DB 2974 CTGACACGCTGGCTGTGGTTTTCTTTCACCTTTCGCTTATGGGTAACTCATGGCTCAGC 3033
QY 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
DB 3034 CACTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGCGAGGTC 3093
QY 940 MetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeuThrAla 959
DB 3094 ATGAGAGAACCATCTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTCATCCGCC 3153
QY 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
DB 3154 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCAGGAGCAGACACCTTGAAGAGCCTGGAT 3213
QY 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
DB 3214 CTCACGAGCAATGCCCTGGGTGACGCTGGGGTTGCTGCACCTGTGCGAGGACCTGAAGCAA 3273
QY 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
DB 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGTGATTTCTGATTTGCTGT 3333
QY 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
DB 3334 GAGGCACTCTCTTGGCCCTTTCCTGCAACCGGCATCTGACCACTGTAAACCTGGTGCAG 3393
QY 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
DB 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCTGTCTCCACGCTCT 3453
QY 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079

```
Db 3454 AACTTACAGATAATTGGGCTGTGGAATGGCAGTACCTGTGCAATAAGGAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLeuLysProArgValVallileaspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGACAGTACTCAAGCCCGAGTCGTAATTGACGGTAGTGTGGCATCTTTTGAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588

RESULT 2
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO0248362.
ACCESSION AX478549
VERSION AX478549.1 GI:22217318
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ramkumar,J. and Arvizu,C.
TITLE Embryogenesis associated proteins
JOURNAL Patent: WO 0248362-A 3 20-JUN-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1..3489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474830CB1"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3489
Score: 5756.00 Matches: 1102
Percent Similarity: 95.34% Conservative: 2
Best Local Similarity: 95.16% Mismatches: 0
Query Match: 81.37% Indels: 54
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AX478549 (1-3489)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAAGGAAGAATTTTCAGCAATTCAGGAATTTACTAAAGAAGAATACTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCCACAGTTTGAATCGAGAATGCCAAGTGGAAATGCTCGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCCTTGATGATATTATGGAGCATCGCTGGCCCTGGGCTACGTCATTAGCATCTTTTGA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLys----- 96
Db 241 AACATGAACCTTCGAAACCTCTCGAGAGAGCCAGGATGACATGAAAAAATTCACAGAA 300
Qy 96 ----- 96
Db 301 GATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTCCAGAAAAATAA 360
Qy 96 ----- 96
Db 361 TATGGCATACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAAACAC 420
Qy 97 -----LysIleSerGlnAlaMetGluGlnGly 106
```

```
Db 421 AAGTATGTTGGAAATTCATTCTTTTTCAGAAATTTTCACAGCTATGGAAAGAGGT 480
Qy 107 AlaThrAlaAlaGluThrGluGluGlnGluIleSerGlnAlaMetGluGlnGluGlyAla 126
Db 481 GCCACAGCAGCAGACAGACAGAAAGAAATTTTCACAGCTATGGAACAAGAGGTGCC 540
Qy 127 ThrAlaAlaGluThrGluGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHis 146
Db 541 ACAGCAGCAGACAGACAGAAAGAAACATGGAGGTGACACATGGGACTACAAGAGTCCAC 600
Qy 147 ValMetThrLysPheAlaGluGluGluAspValArgSerPheGluAsnThrAlaAla 166
Db 601 GTGATGACCAAAATTCGTGAGGAGGAGGTGACGTCTAGTGTGTAATTTGAAAAACATGCTGCT 660
Qy 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
Db 661 GACTGGCCGGAAATGCAAAAGTTGGTGTGCTTTTGAATTCAGACCCGGTGGGCTTCCGG 720
Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206
Db 721 CCTCGCACGGTGGTCTCTGCACGGAAGTCAAGAAATTTGGGAAATCGGCTCTAGCCAGAA 780
Qy 207 IleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTGTGTGCTGGCGCAAGGTGGACTCTACAGGAATGTCTCTACGCTTCTTCTTC 840
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCTCGTTAGAGAGATGCAGCGGAAGAGAGAGAGTGTCTACAGAGTTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCCGAGGCTCCGGTGACGGAGATCATGTGCCGACGAGAAAGCGTGTG 960
Qy 267 PheIleIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTCATCATGTACCGTTCGTGATGACCTGGGCTCTGTCTCTCAACANTACACAAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTGAGAAGCAGACCTCCGTTCCACCTCATACGCAAGTCTGCTGAGAAAGT 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTCCTCCCTGAGTCTTCTGATCGTCACCGTCAGAGACGTGGGCAACAGAAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIle 346
Db 1141 TCAGAGGTGTGTCTCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGNACAAAGNATC 1200
Qy 347 HisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGTCTTCGAGCGCGGATTGGTGAGCATCAGAGACACAAAGGTTGCGTGCATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCys 386
Db 1261 ATGAAACAACCGTAGCTGTCTGACACAGTGCAGGTGCCCGCGCTGGGCTCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCTCTGAGCTGAGGACGTGGTGGGGAGAGCGTGCCTCCCTTCAACCAACCGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCTGTGACGCCGCTTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGCGC 1440
Qy 427 CysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCTCAATCTGAGGAAAGAGTTGCTGAAGCGCTTCTGCAAGCGTCTTCTGCGGTATGGCTGGAGGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGlu 466
```

Db 1501 GTGTGGAAATAGGAAGTCTAGTGTTCACGGTGACGACCTCATGGTTCAAGGACTCGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTCTTTACATGAACATCTCTTCCACAGACGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaLeuTyrTyrValLeu 506
Db 1621 TACTACACCTTTCTCCACCTCAGTCTCCAGGACTTCTGTGGCGCTTGTACTAGTGTTA 1680
Qy 507 GluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGGCTTGAANAATCGACCGAGCTCTGCGCTCTGTACGTTGAGAACAAAGAGGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTyrMetLysArgPheLeuPhe 546
Db 1741 ATGGAGCTTAAACAGCGAGGCTTCCATATCCACTCGCTTTTGGATGAAGCGTTTCTGTGTT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGAGCGAAGACGTAAAGAGGCCACTGGAGGTCTCTGCTGGGCTGTCCCGTTCCC 1860
Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAACCTTCTGCACCTGGGTCTCTGTGGGTCTGACGACCTTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606
Db 1921 ACCACCCAGAGACACCTCGACGCTTCCACTGTCTTTTCGAGACTCAGACACAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeu 626
Db 1981 TTTGTTCCTTGGCATTAACAGCTTCCAGAGGTGGCTTCCGATTAACAGAACCTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATAGCATCTTCTCTGCTCCAGCACTGTCCGTATTTGGCGAAATTCGGGTG 2100
Qy 647 AspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAGGGAATCTTCCCAAGAGATGAGTCCGCTGAGGCAATGCTCTGTGGTCCCTTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATAAGACCTCAATTGAGGAGCAGTGGGAAGATTCTGCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCACCCACACCTCGCGCAGCTGGACCTGGGCAGCAGCATCTTGACAGAGCGGGCCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTGTGTGCNAGCTGAGGCATCCCACTTGCAGATACAGACCTTGATGTTTGA 2340
Qy 727 AsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTACCCCTGGTGTGCAGCACCTCTGGAGATCGTCATGCGCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCys 766
Db 2401 CTAAAGATCCCTCAACTTGGGAGGACCCCACTGGAAGGAAGAGATGTAAGGATGCGGTG 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeu 786
Db 2461 GAAGCCTTAAACACCCCAANAATGTTTGTGGAGTCTTTGAGGCTGGATGCTGTGGATTG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTACTTCAAGATCTCCCAAAATCTTACGACCTCCCCAGCGCTGAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGAGCCTGGCAGGAAACAAGGTGACAGACCGAGGAGTAATGCGCTCTCAGTGATGCTTG 2640

Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTGGCGCCCTGCAGAGCTGATACTGGAGGACTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GGTTCGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGCACACCTGTGCTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCACCAACAGCCTGGGGAACGAGGTGTAATCTACTGTGTCATCATGAGCCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGGCTGATGTAATCAGTCCACCTGGACACGGCTGGCTGTGGT 2880
Qy 907 PheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGCACCTTGGCTTATGGGTAACTCATGTGCTGACGACCTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGAAGACAAATGCGTGAAGCTTCTGTGCGAGTCTATGAGAACCATCTTGTTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966
Db 3001 CTCAGGACCTGGAGTTGGTAAAGTGTCACTCACCGCCGGTGTCTGTGAGAGTCTCTCC 3060
Qy 967 CysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGGAGCAGACACCTGAAGAGCTTGGATCTCACGGAACAATGCCCTGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GACGGTGGGGTGTCTGCACTGTGCGAGGACTGAAGCAAAAGAACAGTGTCTGACGAGA 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGTGAAGCATGTGCACTGACTTCTGATTGCTGTGAGGACACTCTCTCTGGCCCTT 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
Db 3241 TCCTGCAACCGGATCTGACAGTCTAAACCTGGTCAGAAATAACTTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeu 1066
Db 3301 ATGATGAAGCTGTGTGCGCCTTTGCGCTGTCCCACTTAACCTTACAGATAATTGGGCTG 3360
Qy 1067 TrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLys 1086
Db 3361 TGGNAATGGCAGTACCTGTGCAANAATAAGGAAGCTGCTGGAGGAGTGCAGCTACTCAAG 3420
Qy 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3421 CCCCAGTCTGAATTCAGCGTAGTTGGCATCTTTTGTGATGAAGATGACCGGTAC 3474

RESULT 3

AX704821 LOCUS AX704821 3926 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 1 from Patent EP1285964.
ACCESSION AX704821

VERSION AX704821.1 GI:29561487

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Weiss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.

AUTHORS

Human mater proteins

TITLE

Patent: EP 1285964-A 1 26-FEB-2003;

JOURNAL

SCHERING AKTIENGESSELLSCHAFT (DE)

```
FEATURES                                     Location/Qualifiers
source                                     1..3926
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 5756.00 Matches: 1102
Percent Similarity: 95.34% Conservative: 2
Best Local Similarity: 95.16% Mismatches: 0
Query Match: 81.37% Indels: 54
DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AK704821 (1-3926)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGACAAGGAAGAATTTACAGACATTCAGGAATTTACTAAAGAAGAAATCTTCAGAAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCACACAGTTTGAATTCAGAAATCGAGAAATGCCACGTGGAATGTCTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCCTTGATGATATTAAGGAGCATCGCTGGCCCTGGGCTAGCTCCATTAGCAATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLys----- 96
Db 241 AACATGAACCTCGAAACCCCTCTCGAGAGAGGACACGGGATGACATGAAAAATTCACCAGAA 300
Qy 96 ----- 96
Db 301 GATCCTGAAGCAACGATGACTGACAAAGGACCAAGCAAGAAAAAGTGCCAGAAAAATAAA 360
Qy 96 ----- 96
Db 361 TATGGCATGACTAGCTTATCTTGGGGGTGCTGACATCTGACTCGAATAATAAACAC 420
Qy 97 -----LysIleSerGlnAlaMetGluGlnGly 106
Db 421 AAGTATGTTGGAATTCATCTCTTTTCAGAAAATTTCACAAGCTATGGAACAAGAGGT 480
Qy 107 AlaThrAlaAlaGluThrGluGluGlnGluLeuSerGlnAlaMetGluGlnGlyAla 126
Db 481 GCCACAGCAGCAGACAGAGAACAAGAAATTTCACAAGCTATGGAACAAGAGGTGCC 540
Qy 127 ThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHis 146
Db 541 ACACGACGACAGACAGAAACACAGGACATGGAGTGCACATGGGACTACAGAGTCAAC 600
Qy 147 ValMetThrLysPheAlaGluGluGluAspValArgArgSerPheGluAsnThrAlaAla 166
Db 601 GTGATGACCAAAATTCGCTGAGGAGAGGAGTGTAGCTGTAGTTTGAACAACACTGTCTGCT 660
Qy 167 AspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArg 186
Db 661 GACTGGCCGGAATTCGAAACAGTTGGCTGGTGTCTTTGATTACAGACCCGGTGGGGCTCCCG 720
Qy 187 ProArgThrValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArg 206
Db 721 CCTCGCAGCGTGGTCTTCGCACGGAAGTCCAGAAATTCGGAATTCGGTCTTAGCCAGAGG 780
Qy 207 IleValLeuCysTrpAlaGlnGlnGlyLeuTyrGlnGlyMetPheSerTyrValPhePhe 226
Db 781 ATCGTGCTGTGCTGGGCGCAAGGTGGACTCTATACCAGGGAATGTTCTCTCTACGCTCTCTTC 840
```

```
Qy 227 LeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSerArg 246
Db 841 CTCCTCCGTTAGAGAGATGCGACGGGAAGAGAGAGAGAGTGTCTCAGAGTTCATCTCCAGG 900
Qy 247 GluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeu 266
Db 901 GAGTGGCCAGACTCCAGAGCTCCGGTGACGAGATCATGTCCCGACAGAAAAGCTGTTG 960
Qy 267 PheIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCys 286
Db 961 TTCATCATTCACGGTTCGATGACCTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGC 1020
Qy 287 LysAspTrpAlaGluLysGlnProPheThrIleLeuArgSerLeuLeuArgLysVal 306
Db 1021 AAAGACTGGGCTCAGAAAGCAGCCTCCGTTACCTCATACGCACTCTGTCTGAGGAAGGTC 1080
Qy 307 LeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLys 326
Db 1081 CTGCTCCCTGAGTCTCTTCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAG 1140
Qy 327 SerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIle 346
Db 1141 TCAGAGCTCGTGTCTCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGAAACAAGAATC 1200
Qy 347 HisLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIle 366
Db 1201 CACTTGTCTCTGAGCGCGGATTTGTGAGCATCAGAAAGACACANGGGTTGCGTGGCATC 1260
Qy 367 MetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLysCys 386
Db 1261 ATGAAACAACCGTGAGTCTTCGACAGTGCAGAGTCCCGCGCTGGGCTCTCTCATCTGC 1320
Qy 387 ValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeu 406
Db 1321 GTGGCCCTGCAGCTGCAGGACGTGTGGGGAGAGCGTCGCCCCCTTCAACAAACGCTC 1380
Qy 407 ThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArg 426
Db 1381 ACAGGCTCGACGCCCTTTTGTGTTTCATCATAGCTCACCCCTCGAGCGGTGTGCCGCGC 1440
Qy 427 CysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGluGly 446
Db 1441 TGTCTCAATCTGGAGGAAGAGTGTCTCTGAAGCGCTTCTGCCGTATGCTGTGGAGGGA 1500
Qy 447 ValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGlu 466
Db 1501 GTGTGGAATAGGAAGTCAAGTGTGTGACGGTGCAGACCTCATGGTTCAAGGACTCGGGGAG 1560
Qy 467 SerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGlu 486
Db 1561 TCTGAGCTCCGTGCTCTGTTTTCATGAACATCCCTTCCAGACAGCCACTGTGAGGAG 1620
Qy 487 TyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeu 506
Db 1621 TACTACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCTTGTACTACGTGTTA 1680
Qy 507 GluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSer 526
Db 1681 GAGGGCTCGAAATCGAGCCAGCTCTCTGCCCTCTGTAGCTTGAGAAAGACAAAGAGGTCC 1740
Qy 527 MetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPhe 546
Db 1741 ATGGAAGCTTAAACAGGACGAGCTTCCATATCCACTCGCTTTGGAATGAAGGCTTCTTGT 1800
Qy 547 GlyLeuValSerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValPro 566
Db 1801 GGCCTCGTGGCAAGACCGTAGAGGCCACTGGAGTCTCTGCTGGGCTGTCCCGTTCCTCC 1860
Qy 567 LeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAla 586
Db 1861 CTGGGGGTGAAGCAGAAAGCTTCTGCACGTGGGTCTCTCTGTGGGTTCAGCAGCCTAATGCC 1920
Qy 587 ThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGlu 606
```

Db 1921 ACCACCCAGGAGACACCTCGACCCCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAG 1980
Qy 607 PheValArgLeuAlaLeuAsnSerPheGlnGluValTrrPLeuProIleAsnGlnAsnLeu 626
Db 1981 TTTGTTCCTTGGCAATTAACAGCTTCCAGAAGTGTGGCTTCGATTAACACAGAACCTG 2040
Qy 627 AspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgVal 646
Db 2041 GACTTGATAGCATCTTCCTCTCGCCTCCAGCAGCTGTCCGTATTTGGCGAAATTCGGGTG 2100
Qy 647 AspValLysGlyIlePheProArgAspGlnSerAlaGluAlaCysProValValProLeu 666
Db 2101 GATGTCAAAGGAGATCTTCCCAAGAGATAGTCCGCTGAGGAGATGTCTGTGTGCTCCTCTA 2160
Qy 667 TrpMetArgAspLysThrLeuIleGluGlnTrrPLeuAspPheCysSerMetLeuGly 686
Db 2161 TGGATCGGGATAGACCTCATTTGAGGAGCAGTGGGAAGATTTCGTCTCCATGCTTGGC 2220
Qy 687 ThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMet 706
Db 2221 ACCCACCCACACCTCGCGCAGCTGGACCTGGGACAGCAGCATCTTGACAGAGCGGCGCATG 2280
Qy 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
Db 2281 AAGACCTCTGTGCGCAAGCTGAGGCATCCACCTGCAAGATACAGACCTCATGTATTAGA 2340
Qy 727 AsnAlaGlnIleThrProGlyValGlnHisLeuTrrArgIleValMetAlaAsnArgAsn 746
Db 2341 AATGCACAGATTACCCCTTGGTGTGAGCAGCTCTGGAGAATCGTCATGCCAACCGTAAC 2400
Qy 747 LeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluLeuAspValArgMetAlaCys 766
Db 2401 CTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGGAAGAGGATGTAAGGATGGCGTGT 2460
Qy 767 GluAlaLeuLysHisProLysCysLeuLeuGlnSerLeuArgLeuAspCysCysGlyLeu 786
Db 2461 GAAGCCTTAAACACCCAAAATGTTTGTGGAGTCTTTGAGGCTGGATTGTGTGGATTG 2520
Qy 787 ThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSer 806
Db 2521 ACCCATGCTGTACTCGAAGATCTCCCAATCTTTACGACCTCCCGCAGGCTGAATCT 2580
Qy 807 LeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeu 826
Db 2581 CTGACCTGGCAGGAAACAAGGTGACACACGAGGAGTAATGCCTCTCAGTGATGCCTTG 2640
Qy 827 ArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThr 846
Db 2641 AGAGTCTCCAGTGGCCCTGCAGAAAGCTGATCTGGAGGAGCTGTGGCATCACAGCCACG 2700
Qy 847 GlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeu 866
Db 2701 GGTGTCCAGAGTCTGGCCTCAGCCCTCGTACAGACCGAGGCTTGACACACCTGTGCCCTA 2760
Qy 867 SerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuPro 886
Db 2761 TCCACACACGCTGGGAGACGAGGTGTAATCTACTGTGTGTCGATCCATGAGGCTTCCC 2820
Qy 887 HisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGly 906
Db 2821 CACTGTAGTCTGCAGAGGCTGATGTGAATCAGTGCCACCTGGACACGCGCTGGCTGTGGT 2880
Qy 907 PheLeuAlaLeuLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsn 926
Db 2881 TTTCTTGACATGGCCCTTATGGGTAACTCATGGCTGACGACCTGAGCCTTAGCATGAAC 2940
Qy 927 ProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHis 946
Db 2941 CCTGTGGAAGACATGGGCTGAGGCTTCTGTGGAGGTCATGAGAGAACCATCTTGTTCAT 3000
Qy 947 LeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSer 966

Db 3001 CTCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCCCGCGTCTGTGAGAGTGTCTCC 3060
Qy 967 CysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGly 986
Db 3061 TGTGTGATCTCGAGGACAGACACCTGAAGAGCCTGGATCTCAGGACAATGCCCTGGGT 3120
Qy 987 AspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArg 1006
Db 3121 GACGGTGGGCTGTGTCACCTGTGCGAGGAGCTGAAGCAAAAGAACAGTGTCTGACGAGA 3180
Qy 1007 LeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeu 1026
Db 3181 CTCGGGTTGAAGGATGTGGACTGTGATTTGCTGTGAGGACATCTCTCTGGCCCTT 3240
Qy 1027 SerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGly 1046
Db 3241 TCTGTCAACCGGATCTGACCAAGCTTAACCTGGTGACAGATAACTTCAGTCCCAAGGA 3300
Qy 1047 MetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeu 1066
Db 3301 ATGATGAAGCTGTGTCGGCCTTTCCTGCTCCACAGTCTAACTTACAGATAATTTGGCTG 3360
Qy 1067 TrpLysTrpGlnTrrProValGlnIleArgLysLeuGluGluValGlnLeuLys 1086
Db 3361 TGGAAATGGCAGTACCTGTGCAAAATAAGGAAGCTGTGAGGAAGTGCAGCTACTCAAG 3420
Qy 1087 ProArgValValIleAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3421 CCCGAGTCTGTAATTGACGGTAGTTGGCANTCTTTTGATGAGATGACCGGTAC 3474

RESULT 4
AY054986 3885 bp mRNA linear PRI 02-APR-2002
LOCUS Homo sapiens maternal-antigen-that-embryos-require protein (MATER)
DEFINITION mRNA, complete cds.
ACCESSION AY054986
VERSION AY054986.1 GI:19882272
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3885)
AUTHORS Tong Z.B., Bondy C.A., Zhou J. and Nelson L.M.
TITLE A human homologue of mouse Mater, a maternal effect gene essential for early embryonic development
JOURNAL Hum. Reprod. 17 (4), 903-911 (2002)
MEDLINE 21922687
PUBMED 11925379
REFERENCE 2 (bases 1 to 3885)
AUTHORS Tong Z.-B., Bondy, C.A. and Nelson, L.M.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD, NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA

FEATURES
source 1. 3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 3885
/chromosome="19"
/gene="MATER"
1. 3603
/gene="MATER"
/note="oocyte-specific protein; similar to Mus musculus Mater encoded by GenBank Accession Numbers AF074018 and AF143559-AF143573"
/product="maternal-antigen-that-embryos-require protein"
/protein_id="AAU15549.1"
/db_xref="GI:19882273"
SSOPCINMEGDKSLTTFSSYGLWCLVYELDKKEEFQTFKELLKKKSSSTTCSIPQFEIE

NANVECLALLIHEYGASLAWATSISIFENNNRLTISEKARDDMKRHS PEDPRATMTD
QGPSKEVPGISQAVQDSATAAEKEQELSOAMEQSGATAAEETEQEIEIQAMEQEQA
TAATEEQGHGSDTWYKSHWMTFAEEEDVRRSFENTAADPMEQTLAFAFSDRWG
FRPRIVLHGSGIGKALAKRIVLCWAQGLYQGMFSYVFFLPVREMQRKSSVTE
FISREWPSQAPVTEINSRPERLAFIIDGFDLGLVNLNDTKLCKDWAERQPPFTLIR
SLLRKVLPPESFLITVTVRDVGTKEIKSEVWSYPRLLVRGISEGRIHLLERGITGEHQ
KTQGLRAIINNRELLDOCVPAVGSGLICVALOLODVGVESVAPFNQTLTGLHAAFAFH
QLTPRGVVRRCLENERLWIKRFCMAVEGVWNEKSVPDGDDLMVQGLSESELALFH
MNILLPUSHCEEYTFPHLSLQDFCAALYVLEGLEIEPALCPYVSKTKRSMELKHA
GFHLSLWMKRFLFGLGVSEVDVRRPVLGLGCPVLGVQKLLHWSLGLQOPNATPV
DTLDAFLCFETQDEKFRVLRALNSFQWMLPFINQNLDIASFSCLQPCPYLRKIRVDV
KGIIPRDEABEACQVPLVMDRDKTLIEQWEDFCSMGLGTHPLRLQLDLGSSILITERAM
KTLCAKLHPHTKOTLMFRNAOITPGVOHLRIVMANRNLSINLGLSTHKEEDVRM
ACEALHPKCLLESILDCCLTHACYLKI SQUITTSPLSKSLSLAGNKVTDQGVPL
SDALRVQCALQLIEDCGITATGCGSALASALVSNRSLTHLCLSNMNPVEGVNLLUC
RSMRLPHCSLQRLMLNCHDITAGCSLALAMGNSWLTHLSLSMNPVEDGVKLLCE
VMRPSCHLODLVLEKHLTAACCESLSVTSRSHLKSJLDLTDNALDGGVAAALCBG
LKQNSVLTRLGLKACQGLTDCCEALSLSNCRHLTSLNLVQNNFSPKGMWKLCSAF
ACPTSNLIQIIGLWKWQYPVQIRKLEEVQLLKPRVVIDSGSWHSFDEDRYWMKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3885
Score: 5741.50 Matches: 1098
Percent Similarity: 96.16% Conservative: 3
Best Local Similarity: 95.90% Mismatches: 3
Query Match: 81.16% Indels: 41
DB: 9 Gaps: 1

US-10-066-521-6 (1-1344) x AY054986 (1-3885)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTTPCysLeuTyrGlu 20
Db |||
Qy 154 ATGGAAGAGACAAATCGCTCACCTTTTCCAGCTACCGGGCTGCAATGGTGTCTATGAG 213
Db |||
Qy 21 LeuAspLysGluGluPheGlnThrPhelysGluLeuLysLysSerSerGluSer 40
Db |||
Qy 214 CTAGACAAGGAAGAAATTCAGACATTCAGGAATTAAGGAAGAAATCTTCAGATCG 273
Db |||
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaValGluCysLeuAlaLeu 60
Db |||
Qy 274 ACCACATGCTCTATTTCACAGTTCGAAATCGAGAATGCCAACGTGGAATGTCTGGCAC 333
Db |||
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrPheSerIleSerIlePheGlu 80
Db |||
Qy 334 CTCCTTGATGATGATATGAGAGCATCGCTGGCCCTGAGCTAGCTCCATAGATCTTTGAA 393
Db |||
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAspMetLysLys----- 97
Db |||
Qy 394 AACATGAACCTCGGAAACCTCTCGGAGNAGGACGGGATGACATGAAGAAGACATTCACCA 453
Db |||
Qy 97 ----- 97
Db 454 GAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGAAAAAAGTGCACGAAT 513
Qy 98 -----IleSer 99
Db 514 TCACAAGCTGTGAACAAGATAGTGCACAGCTGCAGAGACAAAGAACAGGAATTTCA 573
Qy 100 GlnAlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGlnGlnIleSerGln 119
Db 574 CAAGCTATGGACAAAGAAAGGTGCCACAGCAGCAGACAGACAGAGAACAAAGAAATTTCA 633
Qy 120 AlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGlnGlnIleGlyAsp 139
Db 634 GCTATGGAACAAAGAAAGGTGCCACAGCAGCAGACAGACAGAAACAAAGACATGAGGTGAC 693
Qy 140 ThrThrPheTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg 159
Db 694 ACATGGGACACTCAAGAGTACGATGATGACCAAAATTCGCTGAGGAGGAGGATGATCGTCGT 753
Qy 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
Db |||

Db 754 AGTTTTGAAAACACACTGCTGCTGACTGCGCGGAAATGCAAAACGTTGGCTGGTCTTTTGAT 813
Qy 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
Db 814 TCAGACCGGTGGGGCTTCCGGCCCTCGCACGGTGGTCTTGACGCGAAAGTCAGAAATGGG 873
Qy 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGly 219
Db 874 AAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGCGCAAGGTGACTCTACACAGGA 933
Qy 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
Db 934 ATGTTCTCTACCTCTTCTTCTCCCGTTAGAGAGATGCGCGGAAGAGAGAGACAGT 993
Qy 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
Db 994 GTCACAGAGTTCATCTCCAGGAGTGCCACAGACTCCAGGCTCCGGTGCAGGAGATCATG 1053
Qy 260 SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspLeuGlySerValLeu 279
Db 1054 TCCCGACCAAGAAAGGCTGTGTTCATCATTCAGCGTTTCGATGACCTGGGCTCTGCTCT 1113
Qy 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIle 299
Db 1114 AACATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGGCTCCGTTCACTCATATA 1173
Qy 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
Db 1174 CGCAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTGATCGCTCACCGTCAGAGAC 1233
Qy 320 ValGlyThrGluLysLysSerGluValValValSerProArgTyrLeuLeuValArgGly 339
Db 1234 GTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGTCTCCCGTTACCTGTTAGTTAGAGA 1293
Qy 340 IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
Db 1294 ATCTCCGGGGAACAAGAAATCCACTTGTCTCTTGGCGGGGATTGGTGAAGCATCAGAAG 1353
Qy 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgLysLeuLeuLeuAspGlnCysGlnValPro 379
Db 1354 ACAAAAGGTGGCTGCGATCATCAACACCGTGAGCTGCTCGACCATGCCAGGTGCCAGTGC 1413
Qy 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerVal 399
Db 1414 GCCGTGGCTCTCTCATCTGCTGGGCCCTGCACCTGAGACGCTGGTGGGGAGAGCGCT 1473
Qy 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
Db 1474 GCCCTTCAACCAACGCTCACAGGCTGCACGCGCTTTTTCGGTTTCATCAGCTCACC 1533
Qy 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
Db 1534 CCTCGAGGCTGCTGCGGGCTGCTCAATCTCGAGGAAGAGTTGTCTCTGAAGCGCTTC 1593
Qy 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
Db 1594 TGGCTATGCTGTGAGGAGGTGTGGAATAGGAAGTCAAGTGTGTATGTGACGACCTC 1653
Qy 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
Db 1654 ATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTTCACATGAACATCTTCTC 1713
Qy 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
Db 1714 CCAGACAGCCACTGTGAGGAGTACTACCTCTCCACCTCTCCAGTCTCCAGGACTTCTGT 1773
Qy 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
Db 1774 GCCCTTGTACTACTGTGTAGAGGCTGGAATCGAGCAGCTCTCTGCTCCCTCTGTAC 1833
Qy 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
Db 1834 GTTGAGAAGACAAAGAGGTCCATGAGCTTAAACAGGAGGCTTCCATATCCACTCGCTT 1893

```
Qy 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTTCTCTGTTGGGCTCGTGAGCGAAGACGTAAGAGGCGCACTGGAGTC 1953
Qy 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGGTCCTCTCTG 2013
Qy 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGGTGAGAGCTTAATGCCACACCCAGAGACACCCCTGGAGCGCTTCCACTGCTT 2073
Qy 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
Db 2074 TTCGAGACTCAAGACAAAGATTTGTCCTGGCATTTAAACAGCTTCCCAAGAGTGTGG 2133
Qy 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAACCAAGACCTGGACTTGATAGCATCTTCTCTGCTCCAGCACTGTCCG 2193
Qy 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
Db 2194 TATTGGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGATGAGTCCGCTGAG 2253
Qy 660 AlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGlu 679
Db 2254 GCATGTCTGTGTCCTCTATGATGCGGATTAAGACCTCATTTGAGGAGCAGTGGAA 2313
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCGTCTCCATTCGTGGCACCCACCCACACCTGGCGGACGTGGACCTGGGCGAGC 2373
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTTGACAGCGGGCATGAGACCTGTGTGCCAAGCTGAGGATCCACCTCGCAAG 2433
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGTGTTAGAAATGCACAGATTACCTCCCTGCTGTGCAACACCTCTCGAGA 2493
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 2494 ATCTGTCAATGGCCAAACCGCTTAAGATCCCTCAACTTGGGAGGACCCACCTGGAAGAA 2553
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGTAGGATGGCTGTGAAGCCTTAAACACCCAAATGTTTGTGGAGTCTTTG 2613
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGGATTGCTGTGATTGACCATGCTGCTTTACCTGAAGATCTCCCAATCTTTACG 2673
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2674 ACCTCCCGCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACAGCACCCAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 2734 ACGCTCTCAGTATGCTTGAGGCTTCCAGTGGCCCTTGCAGAGAGCTGATACTGGAG 2793
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCACAGCACCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 2853
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACCTGTGCCTATCAACAACAGCCTGGGGAACGAAGGTGTAATCTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCACTGATGCTGAGAGGCTGATGCTGAATCAGTGCAC 2973
```

```
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGGACACGGCTGGCTGTGTCTCTTTGCACTTATGGCTTATGGGTAACCTCATGGCTGACG 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 3034 CACTGAGCCTTAGCATGAACCCCTGTGGAAGACAAATGGCGTGAAGCTTCTGTGCGAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 3094 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 GCGTGTCTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 3214 CTCAGGACAATGCCCTGGGTGACGGTGGGTTGCTGCGCTGTCCGAGGACCTGAAGCAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 3274 AAGACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGACTGACTTCTGATTGCTGT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 3334 GAGGCACTCTCTGGCCCTTCTCTGCAACCGGCATCTGACCAGTCTAAACCTGGTGCAG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCCTTGTCCCTGTCCACGCTCT 3453
Qy 1060 AsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATAATTGGGCTGTGGAATGTCAGTACCCCTGTGCAATGAAGAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAAGTGCGAGCTACTCAAGCCCCGAGTGTGTAATTGACGGTAGTTGGCATTTCTTTGAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
RESULT 5
LOCUS AX427610 3900 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 23 from Patent WO0232955.
ACCESSION AX427610
VERSION AX427610.1 GI:21537730
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source 1. .3900
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1. .3603
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD35973.1"
/db_xref="GI:21537731"
/db_xref="RENTREMBL:CAD35973"
/translation="MKVAGGLELGAALIASPRALVTLTSTGPTCSILPKNPLFPQNL
SSQPCIKMGDKSLTSPSSYGLQWCLYELDKKEEFQTKELLKKKSBSSTTCSIPQPIE
```


NANVECLALLHYYVGASLAWATTSISIFENNNLRTLSEKARDMDKRSHPEDPRATWTD
 QGSEKVPGISQAVQDSATAETKEQEI SQAMEQDGAATAETEEQEI SQAMEQEGA
 TAATEHQGGDGTWKASHVMTKFABEEDVRKSFENTAADNPEMQTLAFAFSDRWG
 FRPTVVLHGKSGIGKSALARRIVLCWAQGLGYQMFPSYFFFLPFRMORKESSVTE
 FISREWPDQAQVTEIMSRPERLLFIIDGDDLGSLVNDNTKLCOWAEKQPPFTLIR
 SLARKVLLPESFLIVTVRWGTEKLVKEVSPRYLLVRGISGEORHLHLLERIGIRBHO
 KTOGLRALINRELLDQOVPAVGSLLICVALQLODVVGESVAPPNQTLTGLHAAPAPH
 QLTFRGVRRCLNLEERVLKFRFCMAVEGVWNRKSVFDDGDLMVQGLGESELRALUFH
 MNILLPSHCHEEYTFPHLSQDFCAALYYVLEGLEIEPALCLPLYVEKTRKSMELKQA
 GFHLSLWKRFLFGLVSEVDRPLEVLLGCPVLGVQKLLHWSLLGQOPNATTPG
 DTLDAFCLPETODKEFVRILALANSFOEWLPIINONLDLIASSFCLQHCPCYRKRIRDRAM
 KGIIPRDESAEACPVVLMWRDKTLIEEOMEDFCMLGTHPHRLQDLGLSSILITERAM
 KTLCAKLHPCTKOTLAFNAQITPGVOHLRIVMAFNRLRSINLGGTHLKEEDVRM
 ACEALKPKLLSRLDCCGLTHACYIKISQILTTSPSLKSLAGNKVTDQGVTP
 SDARVSOCLALQILDEDCGITATGCSLASALVSNRSLTHLCSNNSLNNEGVNLC
 RSMRLPCHSQRLMNOCHLDTAGCSLALAMGNSWLTHLSLMPNVEDNGVKLJCE
 VMREPSCHLOLEVLKCHLTAACCESLSVCVSRSHLKSJLDLTDNALGDGVAAALCEG
 LKOKNSVLTIRGLKACGLTSDCEALSLSLSCNRHLTSLNLVONNFSPKGMMLKCSAF
 ACPTSNLIQIIGLKWQYVPQIRKLLBEEVQLLKPRVVIDGSHWSHDEDDRYWKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3900
 Score: 5741.50 Matches: 1098
 Percent Similarity: 96.16% Conservative: 3
 Best Local Similarity: 95.90% Mismatches: 41
 Query Match: 81.16% Indels: 1
 DB: 6 Gaps: 1

US-10-066-521-6 (1-1344) x AX427610 (1-3900)

QY 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTTPCysLeuTyrGlu 20
 DB 154 ATGAAGAGAGACAAATCGCTACCTTTTCAGACTACGGGCTGCAATGGTGTCTCTATGAG 213
 QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysSerSerGluSer 40
 DB 214 CTAGACAGAGAGAAATTTACAGATTTCAAGGAATTTACTAAGAAGAAATTTTCAGAAATCG 273
 QY 41 ThrThrCysSerIleProGlnPheGluIleGluAlaSerAlaValGluCysLeuAlaLeu 60
 DB 274 ACCACATGCTCTATTCCACAGTTTGAATCGAGAAATCGAGAAATGCTTGGCACTC 333
 QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaThrAlaThrSerIleSerIlePheGlu 80
 DB 334 CTCTTGATGAGTATTATGAGCATCGCTGGCTGGCTACGTCATTAGCATCTTTGAA 393
 QY 81 AsnMetAsnLeuAArgThrLeuSerGluLysAlaAArgAspAspMetLysLys----- 97
 DB 394 AACATGAACCTGCGAACCCCTCTCGAGAGAGGCACGGATGACATGAAGAACATTCACCA 453
 QY 97 ----- 97
 DB 454 GAAGATCTCTGAAGCAACGATGACTGACCAAGACCACCAAGCAAGGAAAAAGTGCCAGGAATT 513
 QY 98 -----IleSer 99
 DB 514 TCACAAGCTGTGCAACAGATAGTGCCACAGCTGCAGAGACAAAGAACAGGAAATTTCA 573
 QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGluIleSerGln 119
 DB 574 CAAGCTATGGNACAAAGAGGTGTCACAGCAGCAGCAGACAGACAGAACAGAAATTTCAAA 633
 QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAsp 139
 DB 634 GCTATGGAACAAGAGGTGTCACAGCAGCAGCAGACAGACAGAACAAAGACATGAGGTGAC 693
 QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValAArgArg 159
 DB 694 ACATGGGACATACAAAGTTCAGTATGACCAAAATTCGCTGAGGAGGAGGTACGTCGT 753
 QY 160 SerPheGluAsnThrAlaAlaAspTyrProGluMetGlnThrLeuAlaGlyAlaPheAsp 179

DB 754 AGTTTTTGAACACACTGCTGCTGACTGGCCGGAATAATGCAAAACGTTGGCTGGTCTTTGAT 813
 QY 180 SerAspArgTyrGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
 DB 814 TCAGACCGGTGGGGCTTCGGCCCTCGCACGGTGGTCTTGACGCGAAAGTCAGGAATTGGG 873
 QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTyrAlaGlnGlyLeuTyrGlnGly 219
 DB 874 AAATCGGCTCTACCCAGAGAGGATCGTGTGTCGGCGCGCAAGTCGACTCTTACCAGGGA 933
 QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
 DB 934 ATGTTTCTCTACGTCCTTCTTCTCCCGTTAGAGAGATGCGAGCGGAGAGAGAGAGCAGT 993
 QY 240 ValThrGluPheIleSerArgGluTyrProAspSerGlnAlaProValThrGluIleMet 259
 DB 994 GTCACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGGAGATCATG 1053
 QY 260 SerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAspLeuGlySerValLeu 279
 DB 1054 TCCCGACCAAGAAAGGCTGTTGTTTCATCATGACGGTTCGATGACCTGGGCTCTGTCCTC 1113
 QY 280 AsnAsnAspThrLysLeuCysLysAspTyrAlaGluLysGlnProProPheThrLeuLeu 299
 DB 1114 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGCAGGCTCGTTTCAACCTCAT 1173
 QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
 DB 1174 CGCAGTCTGCTGAGGAAGGTCTGCTCCGTGAGTCTCTCTGATCTCTCCGTCAGAGAC 1233
 QY 320 ValGlyThrGluLysLysSerGluValValValSerProArgTyrLeuLeuValArgGly 339
 DB 1234 GTGGGACACAGAGAGCTCAAGTCAGAGGTCTGTCCTCCCGTTTACCTGTGTAGTAGGA 1293
 QY 340 IleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
 DB 1294 ATCTCGGGGACAAAGAAATCCATTTGCTCTTGAGCGGGGATTTGGTGGACATCAGAAG 1353
 QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValPro 379
 DB 1354 ACACAAAGGTTGGTCGATCATCAACAAACCGTGAGCTGCTCGACCAAGTCCAGAGTGCCCC 1413
 QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGluSerVal 399
 DB 1414 GCGTGGGCTCTCTCATCTGCGGGCCCTGCACTGCAAGCTGAGAGCTGGTGGGGAGAGCGTC 1473
 QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
 DB 1474 GCGCCCTTCAACCAACGCTCACAGGCTGCAGCGCGCTTTTGGCGTTTCATCAGCTCACC 1533
 QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
 DB 1534 CCTCGAGGGCTGTCGGGCTGCTCTCAATCTGGAGGAAGAGTTGCTCTGAAAGCGCTTC 1593
 QY 440 CysArgMetAlaValGluGlyValTyrAsnArgLysSerValPheAspGlyAspAspLeu 459
 DB 1594 TGGCGTATGCTGTGAGGAGGTGCGAATAGGAAGTCAGTGTGTGATGTGACGACCTC 1653
 QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
 DB 1654 ATGTTTCAAGGACTCGGGAGTCTGAGCTCGGTGCTCTGTTTTCACATGAACATCTTCTC 1713
 QY 480 ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
 DB 1714 CCAGACAGGCACCTGTGAGGAGTACTACACTTCTCCACTCTCCAGTCTCCAGGACTTCTGT 1773
 QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
 DB 1774 GCGGCTTGTACTACGTGTGAGGGCCCTGGAAATCGAGCAGCTCTCTGCCCTCTGTGAC 1833
 QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
 DB 1834 GTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCCACTCGCTT 1893

Qy 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeuGluVal 559
Db 1894 TGGATGAAGCGTTTCTCTGTTGGCCCTCGTGAGCGAAGACGTAAGAGGCGCACTGAGGTC 1953
Qy 560 LeuLeuGlyCysProValProLeuGlyVallyGlnLysLeuLeuHisTrpValSerLeu 579
Db 1954 CTGCTGGGCTGTGCCCTTCCCTGGGGTGAAGCAGAAAGCTTCTGCACATGGGCTCTCTG 2013
Qy 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
Db 2014 TTGGTCAAGACCTTAATGCCCACACCCAGAGACACCTTGGAGCCCTTCCACTGTCTT 2073
Qy 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAnSerPheGlnGluValTrp 619
Db 2074 TTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCCAAGAGTGTGG 2133
Qy 620 LeuProLleAnGlnAnSerLeuLleAlaSerSerPheCysLeuGlnHisCysPro 639
Db 2134 CTTCCGATTAACCAAGACCTGGACTTGTATAGCATCTTCTCTGCTCCAGCACTGTCCG 2193
Qy 640 TyrLeuArgLysIleArgValAspVallyGlnLysPheProArgAspGluSerAlaGlu 659
Db 2194 TATTTGCGGAATTCGGGTGATGTCAAAGGGATCTTCCCAAGAGATGATCCGCTGAG 2253
Qy 660 AlaCysProValProLeuTrpMetArgAspLysThrLeuLleGluGluGlnTrpGlu 679
Db 2254 GCATGTCTGTGGTCCCTATGATGCGGATAGACCTCTCATGAGGACAGTGGGA 2313
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 2314 GATTTCTGCTCATCTTGGCACCCACCCACACCTTGGCGGAGCTGGACCTGGGCGAGC 2373
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGACATCCCACTGCAAG 2433
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 2434 ATACAGACCTGATGTTAGAAATGCACAGATTACCTGCTGTGTGCAACACTCTGGAGA 2493
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAnLeuGlyThrHisLeuLysGlu 759
Db 2494 ATCGTCATGGCGCAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCGCCACCTGAAGGA 2553
Qy 760 GluAspValArgMetAlaCysGluAlaLysHisProLysCysLeuLeuGluSerLeu 779
Db 2554 GAGGATGAAGATGGCGTGTGAAGCCCTTAAACACACCCCAAAATGTTGTGGAGTCTTTG 2613
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 2614 AGGCTGGATTTGCTGGATTGACCATGCTGTACCTGAAGATCTCCCAATCTCTACG 2673
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 2674 ACCTCCCCAGCCTGAATCTCTGAGCTGGCAGGAAACAAGTGCACAGACAGGAGTA 2733
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuLleLeuGlu 839
Db 2734 ACGCTCTCAGTGATGCTTGGGGTCTCCAGTGGCCCTTGCAGAAAGCTGATATCTGGAG 2793
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 2794 GACTGTGGCATCACAGCACCGGTTGCCAGAGTCTGGCTCAGCCCTCTGTCAGCACCGG 2853
Qy 860 SerLeuThrHisLeuCysLeuSerAnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 2854 AGCTTGACACACTGTGCTATCCCAACAACAGCTGGGGAACGAAGGTGTAATCTACTG 2913
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAnGlnCysHis 899
Db 2914 TGTGATCCATGAGGCTTCCCACTGTATGTGACAGAGGCTGATGCTGAATCAGTGGCCAC 2973

Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 2974 CTGACACGCTGGCTGTGGTCTCTTGTGCACTTATGGCTTATGGTAACCTCATGGCTCAGC 3033
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallyLysLeuLeuGluVal 939
Db 3034 CACTGAGCCTTAGCATGAACCCCTGTGGAAGACAATGGCTGAAGCTTCTGTGCGAGGTC 3093
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuVallyLysCysHisLeuThrAla 959
Db 3094 ATGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACGCC 3153
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 3154 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGAT 3213
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 3214 CTCAGGACAATGCCCTGGGTGACGGTGGGTGTGCTGCTGCGAGGAGCTGAAGCAA 3273
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCys 1019
Db 3274 AGAACAGTGTCTGACGACACTCGGTTGAGGCGATGTGACTGACTTCTGATTGTCT 3333
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAnLeuValGln 1039
Db 3334 GAGGCACTCTCTGGCCCTTCTGCAACCGGCATCTGACCACTGTAAACCTGGTGCAG 3393
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 3394 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGCGCCTTGTCTGCTCCACGCT 3453
Qy 1060 AsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLysLeuLeu 1079
Db 3454 AACTTACAGATTAATGGGCTGTGGAATGCGACTACCTGTGCAATAGGAAGCTGCTG 3513
Qy 1080 GluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSerPheAsp 1099
Db 3514 GAGGAGTGCAGCTACTCAAGCCCGAGTGTAAATTCAGCGTAGTTGGCATTTCTTTGAT 3573
Qy 1100 GluAspAspArgHis 1104
Db 3574 GAAGATGACCGGTAC 3588
RESULT 6
AX459881 LOCUS AX459881 5859 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 24 from Patent WO0240668.
ACCESSION AX459881
VERSION AX459881.1 GI:21725649
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tschoopp, J. and Martinon, F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 24 23-MAY-2002;
FEATURES Apotech Research and Development Ltd. (CH)
source Location/Qualifiers
1. 5859
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP8/Py12---Py8.dna"
ORIGIN
Alignment Scores: 0 Length: 5859
Pred. No.: 5724.50 Matches: 1106
Score: 90.53%
Percent Similarity: 90.53%

Best Local Similarity:	90.29%	Mismatches:	7
Query Match:	80.92%	Indels:	109
DB:	6	Gaps:	3
US-10-066-521-6 (1-1344) x AX459881 (1-5859)			
QY	1	MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu	20
DB	1975	ATGGAAGAGAGCAAAATCGCTCACCTTTTCCAGCTACCGGCTGCAATGGTGTCTATGAG	2034
QY	21	LeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLysLysSerSerGluSer	40
DB	2035	CTAGACAAGGAAGAAATTCACACATTCAAGGAATTAATAAGAGAANAATCTTCAGAAATCG	2094
QY	41	ThrThrCysSerIleProGlnPheGluIleGluAenAlaenValGluCysLeuAlaLeu	60
DB	2095	ACCACATGCTCTATTTCACACAGTTTGAATTCAGAAATCCAGAAATGCTCGGCACATC	2154
QY	61	LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu	80
DB	2155	CTCTTGCAATGAGTAATATGGAAGCATCGCTGGCCCTGGGCTACGTCATTAGCATCTTTGAA	2214
QY	81	AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys-----	97
DB	2215	AACATGAACCTGCGAACCTCTCGAGAGGACCGGGATGACATGAAAGACATTACCA	2274
QY	97	-----	97
DB	2275	GAAGATCCTGAAGCAACGATGACTGACCAGGACCAAGCAAGGAAAAAATGCCAGGAATT	2334
QY	98	-----IleSer	99
DB	2335	TCACAAGCTGTGCACAAGATAGTGCACAGCTCGACAGACAAAGAACAAAGAAATTTCA	2394
QY	100	GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln	119
DB	2395	CAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAGAGAACAAGAAATTTCAAA	2454
QY	120	AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyAsp	139
DB	2455	GCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAGAAACAAGACATGGAGGTGAC	2514
QY	140	ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArg	159
DB	2515	ACATGGGACTCAAGAGTTCAGTGATGACCAAAATTCGCTGAGGAGGAGATGATACGTCGT	2574
QY	160	SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp	179
DB	2575	AGTTTGAACACACTGCTGCTGACTGCGCGAAATGCAAAACGTTGGCTGCTTTTGAT	2634
QY	180	SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly	199
DB	2635	TCAGACCGGTGGGCTTCCGCGCTCGCACGGTGGTTCTGACGCGAAAGTCAGGAATTTGGG	2694
QY	200	LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGly	219
DB	2695	AAATCGGCTTAGCCAGAAAGATCGTCTGTGCTGGGCGCAAGTGACTCTACAGGGA	2754
QY	220	MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer	239
DB	2755	ATGTTCTCTACGCTTCTTCTCTCCCGTTAGACAGATGCAGCGGAAGAGGAGAGCAGT	2814
QY	240	ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet	259
DB	2815	GTCACAGATTCACTCTCCAGGAGTGGCCAGACTCCCAGGCTCCGCTGACCGAGATCATG	2874
QY	260	SerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySerValLeu	279
DB	2875	TCCCGACCAAGAAAGCTGTGTTTCAATGACCGGTTTCGATGACCTGGGCTGTGCTCTC	2934
QY	280	AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle	299
DB	2935	AACAAATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAAGCAGCGCTCGCTTCAACCTCATA	2994

QY	300	ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuLeuValThrValArgAsp	319
DB	2995	CGCAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTGATCGTCAACCGTCAGAGAC	3054
QY	320	ValGlyThrCluLysLeuLysSerGluValValValSerProArgTyrLeuLeuValArgGly	339
DB	3055	GTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGTCTCCCGTTCACCTGTAGTTAGAGGA	3114
QY	340	IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys	359
DB	3115	ATCTCCGGGGAACAAGAATCCACTTGTCTCTTGAGCGGGATTTGGTGAGCATCAGAAG	3174
QY	360	ThrGlnGlyLeuArgAlaIleMetAsnAsnArgLysLeuLeuAspGlnCysGlnValPro	379
DB	3175	ACAAAGGGTTGCTGCGATCATGAACAACCGTGAGCTGCTCGACCAAGTCCAGAGTGCCC	3234
QY	380	AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal	399
DB	3235	GCCGTGGGCTCTCTCATCTGCTGGGCCCTGCAGCTGCAGACGTTGGTGGGGAGAGCGTC	3294
QY	400	AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr	419
DB	3295	GCCCCCTTCAACCAACAGCTCACAGGCTGCACGCGCTTTTGTGTTCATCAGCTCACC	3354
QY	420	ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe	439
DB	3355	CCTCGAGGCGTGTCTCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTCTGAAGCGCTTC	3414
QY	440	CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspAspLeu	459
DB	3415	TGCCGTATGCTGTGAGGAGGTGTGGAATAGGAAGTCAAGTGTTCACGCTGACGACCTC	3474
QY	460	MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu	479
DB	3475	ATGTTTCAAGGACTCGGGAGTCTGAGCTCGGTGCTCTGTTTCACATGAACATCTTCTC	3534
QY	480	ProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCys	499
DB	3535	CCAGACGACCTGTGAGGAGTACTACACTTCTCCACTCTCAGCTCTCCAGGACTTCTGT	3594
QY	500	AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr	519
DB	3595	GCGGCTTGTACTACTGTGTAGAGGCTTGAAATCGAGCCAGCTCTCTGCCCCCTCTGTAC	3654
QY	520	ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu	539
DB	3655	GTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGGCTTCCATATCCACTCGCTT	3714
QY	540	TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal	559
DB	3715	TGGATGAACGTTTCTGTTTGGCTCGTAGCGAAGACGTAAAGAGGCGCACCTGGAAGGTC	3774
QY	560	LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu	579
DB	3775	CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGACCTGGGTCTCTCTG	3834
QY	580	LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu	599
DB	3835	TTGGGTCAGCAGCTTAATGCCACACCCAGGAGACACCTTGGAGCGCTTCCACTGCTCTT	3894
QY	600	PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp	619
DB	3895	TTCCAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCAAGAGGTGG	3954
QY	620	LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro	639
DB	3955	CTTCCGATTAACACAGAACTGGACTTGATGACATCTTCTCTGCTCCAGCACTGTCCG	4014
QY	640	TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu	659
DB	4015	TATTTGGGAAAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGATGTCGCTGAG	4074

```
Qy 660 AlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
Db 4075 GCATGCTCTGGTGGTCCCTCTATGATGGGATGAAGACCCCTCATTTGAGGACAGTGGAA 4134
Qy 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
Db 4135 GATTTCTGCTCCATGCTTGGCACCCACCCACCTCGCGGAGCTGGACCTGGGAGCAGC 4194
Qy 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719
Db 4195 ATCTGCAGAGAGCGGCCCATGAGACCTGTGTGCCAGCTGAGGCATCCACCTGCAG 4254
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 4255 ATACAGACCTCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGCACCTCTCGAGA 4314
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLysLysGlu 759
Db 4315 ATCGTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGAGCCACCTGAAGAA 4374
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 4375 GAGGATGTAAAGATGGCGTGTGAAGCCTTAACAACCCCAAAATGTTTGTGGAGTCTTTG 4434
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 4435 AGGCTGGATTGCTGTGGATTGACCATGCTGTACCTGGAAGATCTCCCAATCCTTAACG 4494
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 4495 ACCTCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGCACCGAGTA 4554
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 4555 ATGCCTCTCAGTGATGCTTCAGAGTCTCCAGTGGCGCCCTGCAGAGCTGATACTGGAG 4614
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 4615 GACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 4674
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 4675 AGCTTGACACACCTGTGCTTATCAACAACACCGCTGGGGAAACGAAGGTGTAATCTACTG 4734
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 4735 TGTGATCCATGAGCTTCCCACTGTAGTCTGCAGAGCTGATGCTGAATCAGTGCAC 4794
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 4795 CTGGACACGGCTGGCTGTGGTTTCTTTCACCTTGGCTTATGGGTAACTCATGCTGACG 4854
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 4855 CACCTGAGCCTTAGCATGAACCCCTGTGAAGACAAATGGCGTGAAGCTTCTGTGCGAGTC 4914
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 4915 ATGAGAAACCATCTTGTCTATCTCAGAGCTGGAGTTGGTAAAGTGTCACTCACCGCC 4974
Qy 960 AlaCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 4975 GCGTGTCTGAGAGTCTCTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCTGGAT 5034
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 5035 CTCACGGCAATGCCCTGGGTGACGGTGGGGTGTGTCACCTGTGCGAGGAGTCAAGCAA 5094
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 5095 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGTGAGCTGACTTCTGATGTGCTGT 5154
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
```

```
Db 5155 GAGGCATCTCTCTTGGCCCTTTCTCTCAACCGGCATCTGACCAGTCTAAACCTGGTGCGAG 5214
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 5215 AATAACTTCAGTCCCAAGAGGAATGATGAAGCTGTGTTGGGCCCTTTGCTGTGCCACGTCT 5274
Qy 1060 AsnLeuGlnIleIleGly----- 1065
Db 5275 AACTTACAGATANTTGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAATCTGCTCTA 5334
Qy 1065 ----- 1065
Db 5335 GTAGTTTTTCTTAAAGTCACTGTTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 5394
Qy 1065 ----- 1065
Db 5395 CTGTACCTCCCCAAACCCCTATAACAGACCCCGGACAGAGACTCTGAGGAAGAAATTCAT 5454
Qy 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 5455 GGATGGACTGAAGGCTGTGGAATGGCACTACCTGTGCAATNAGGAAGCTGCTGGAG 5514
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 5515 GAAAGTGCACTACTCAAGCCCGAGTGTAAATGACGGTAGTTGGCAATCTTTTGATGAA 5574
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 5575 GATGACCGATTTGATTTCAAAGCCAGCAAAACAGTCACCTCAGCAAGACAGACTTACAAT 5634
Qy 1112 LeuProGluSerArg 1116
Db 5635 CTATGGCATCAGAG 5649
RESULT 7
AX459873 LOCUS AX459873 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 16 from Patent WO0240668.
ACCESSION AX459873
VERSION AX459873.1 GI:21725645
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Tschoopp,J. and Martinon,F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 16 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source 1. 6939
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP5/Py8.dna /GENSCAN_predicted_CDS_1/ 6939_bp"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
US-10-066-521-6 (1-1344) x AX459873 (1-6939)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTrpGlyLeuGlnTrpCysLeuThrGlu 20
Db 3055 ATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 3114
```

QY 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
DB 3115 CTAGACAAGGAAGAAATTTTCAGACATTTCAAGAAATTTACTAAAGAAAGAAATTTTCAGAAATCG 3174
QY 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
DB 3175 ACCACATGCTCTATTCCACAGTTTGAATCGAGATGCCACGTGGAAATGTCCTGGCACTC 3234
QY 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
DB 3235 CTCCTTGATGATGATTATGAGACATCGCTGGCCCTGGCTACGTCATAGCATCTTTGAA 3294
QY 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLys----- 97
DB 3295 AACATGAACCTTGCAGAACCTCTCGAGAAAGCAGCGGATGACATGAAGAAAGACATTCACCA 3354
QY 97 ----- 97
DB 3355 GAAGATCTCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAGTGCACAGGAATT 3414
QY 98 -----IleSer 99
DB 3415 TCACAAGCTGTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAAATTTCA 3474
QY 100 GlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGln 119
DB 3475 CAAGCTATGGAAACAAGAGGTGCCACAGCAGCAGACAGAGACAGAAAGAAATTTCAAA 3534
QY 120 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGlyLys 139
DB 3535 GCTATGGNACAAAGAGGTGCCACAGCAGCAGCAGACAGACAGACAGACATGGAGGTGAC 3594
QY 140 ThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspValArgArg 159
DB 3595 ACATGGGACTACAAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGTAGCTCGT 3654
QY 160 SerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAsp 179
DB 3655 AGTTTTGAATAACACTGCTGCTGACTGGCCGGAATGCAAACTGGCTGGCTTTTGAT 3714
QY 180 SerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGly 199
DB 3715 TCAGACCGGTGGGCTTCGCGCTCGCACGCTGTTCTGCAAGAAAGTCAGGAAATTTGGG 3774
QY 200 LysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyrGlnGly 219
DB 3775 AAATCGGCTTAGCCAGAAAGGATCGTGTGCTGGCGCGCAAGGTGGACTCTACCAAGGA 3834
QY 220 MetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSer 239
DB 3835 ATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAGGAGAGCAGT 3894
QY 240 ValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMet 259
DB 3895 GTCACAGATTCACTCCAGGGAGTGGCCAGACTCCAGGCTCCGFTGACGGAGATCATG 3954
QY 260 SerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySerValLeu 279
DB 3955 TCCCGACCAAGAGCGTGTTCATCATTCAGCGTTTCGATGACCTGGGCTCTGTCTC 4014
QY 280 AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThrLeuIle 299
DB 4015 AACAAATGACAAAGAGCTCGCAAGACTGGGCTGAGAAAGCAGCGCTCCGTTCAACCCCTCATA 4074
QY 300 ArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAsp 319
DB 4075 CGCAGTCTGCTGAGAAAGTCTGCTCCCTGAGTCTTCTGATGCTCACCGTCAGAGAC 4134
QY 320 ValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGly 339
DB 4135 GTGGGCACAGAGAGCTCAAGTCAGAGGTCGTGTCTCCCCGTTACCTGTTAGTAGGA 4194

QY 340 IleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLys 359
DB 4195 ATCTCCGGGGAAACAAAGAAATCCACTTGCTCTTGAGCGGGGATTTGGTAGCATCAGAAG 4254
QY 360 ThrGlnGlyLeuArgAlaIleMetAsnAsnArgLeuLeuLeuAspGlnCysGlnValPro 379
DB 4255 ACACAGGGTGTGGTGATCATGAACACCGTAGCTGCTCAGCAGTGCCAGGTGCCCC 4314
QY 380 AlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerVal 399
DB 4315 GCGTGGGCTCTCTCATCTGCGTGGCCTGCAGCTGCAGGACGTGTGTGGGGAGAGCGTC 4374
QY 400 AlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThr 419
DB 4375 GCCCCTTCAACCAACAGCTCACAGGCGCTGCACGCGCTTTTGTGTTTCATCAGCTCACC 4434
QY 420 ProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPhe 439
DB 4435 CCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCTGAAGCGCTTC 4494
QY 440 CysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeu 459
DB 4495 TGCCGTATGGCTGTGAGGAGGTGTGGAATAGGAAGTTCAGTGTTCACGCTGACGACCTC 4554
QY 460 MetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeu 479
DB 4555 ATGTTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGAACATCTTCTC 4614
QY 480 ProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAspPheCys 499
DB 4615 CCAGACAGCCACTGTGAGGAGTACTACACTTCTCCACTCAGTCTCCAGGACTCTCTGT 4674
QY 500 AlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyr 519
DB 4675 GCGCCTTGTACTACTGTTAGAGGCGCTGGAATCGAGCCAGCTCTCTGCCCTCTGTAC 4734
QY 520 ValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeu 539
DB 4735 GTTGAAGAACAAAGAGGTCCATGGAGCTTAAACAGGCGAGGCTTCCATATCCACTCGCTT 4794
QY 540 TrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeuGluVal 559
DB 4795 TGGATGAAGGTTTCTTGTGTCCTCGTAGCGAAGAGCTAAGAGGCGCACCTGGAGGTC 4854
QY 560 LeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeu 579
DB 4855 CTGCTGGGCTGTCCCTTCCCTGGGGGTGAAGCAGAAAGCTTCTGCACTGGGTCTCTCTG 4914
QY 580 LeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeu 599
DB 4915 TTGGGTGACGAGCGCTAATGCCACCCAGGAGACACCTCGAGCGCTTCCACTGCTCTT 4974
QY 600 PheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrp 619
DB 4975 TTTCGAGACTCAACACAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAGAAAGTGTGG 5034
QY 620 LeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysPro 639
DB 5035 CTTCCGATTAAACAGAAACCTGGACTTGTATAGCATCTTCTTCTGCTCCAGCACTGTCCG 5094
QY 640 TyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGlu 659
DB 5095 TATTTCCGAAAAATTCGGGTGATGTCAAAGGGATCTTCCCAAGAGATAGTCCGCTGAG 5154
QY 660 AlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGlu 679
DB 5155 GCATGTCCTGTGGTCCCTCTATGGATCGGGATAGACCTCTATTGAGAGAGCAGTGGGA 5214
QY 680 AspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSer 699
DB 5215 GATTTCTGCTCCATGTTGGCACCCACCACTGCGGCAGCTGGACCTGGGCGCAGCAGC 5274
QY 700 IleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLys 719

Db 5275 ATCTTCAGACAGCGCGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAG 5334
Qy 720 IleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArg 739
Db 5335 ATACAGACCTGTAGTTTAAATGACAGATTACCCCTGGTGTGACAGACCTCTGGAGA 5394
Qy 740 IleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGlu 759
Db 5395 ATCGTCATGGCCAAACCGTAACCTTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGAA 5454
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAAGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTGTGGAGCTCTTG 5514
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysValLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATTGCTGGATTGACCCATGCCCTGTACCTGAAGATCTCCCAATCTTACG 5574
Qy 800 ThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCAGCCTGAATCTCTGAGCCTGCGAGAAACAAAGGTGACAGACCCAGGAGTA 5634
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 5635 ATGGCTCTCAGTGATGCTTGAAGTCTCCAGAGTCTCCAGTGGCGCCCTGCAGAAAGCTGATCTGGAG 5694
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGGCANTACAGCCAGCGGTTGCCAGAGTCTGGCCCTCAGCCCTCGTCAGCAACCGG 5754
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 5755 AGCTTGACACACCTGTGCTTATCAACAAACAGCCTGGGGAAGAGTGTGAATCTACTG 5814
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTGATCCATGAGGCTTCCCACTGTAGTGTGACAGAGCTGTGCTGAATCAGTGGCCAC 5874
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGACAGCGCTGGCTGGTGTGTTCTTGCACTTGGCTTATGGGTAACTCATGCTCAGC 5934
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 5935 CACCTGAGCCTTAGCATGAACCTGTGGAGACAATGGCGTGAAGCTTCTGTGCGAGGTC 5994
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACGCC 6054
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACACACCTGAAGAGCCTGGAT 6114
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCAGGACAANTGCCCTGGGTGACGGTGGGTGTGCTGCACCTGTGCGAGGACTGAAGCAA 6174
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGACCTGCTTCTGATTGCTGT 6234
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGACACTCTCTTGGCCCTTCTTCGCAACCGGCATCTGACAGCTGTAAACCTGGTGCAG 6294
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAGCTGTGTTGGGCTTGTGCTGCCACGCTCT 6354
Qy 1060 AsnLeuGlnIleLeuGly----- 1065

Db 6355 AACTTACAGATAAATGGCAATGACTCTGAAGAAAAATGACGTTCTTCGAGAAATCTGCTCTA 6414
Qy 1065 ----- 1065
Db 6415 GTAGTTTGTCTAAAGTCACTGTTTCCAGAACCTATCAATGACATTAAGGGAGAACTTA 6474
Qy 1065 ----- 1065
Db 6475 CTGTACCTCCCAACCCCTATAACAGACCCCGGCACAGACTCTGAGGAAGGAATTCAT 6534
Qy 1066 -----LeuTrpLysTrpGlnTyProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GSATGAGACTGAAGGCTGTGGAATGCGAGTACCCCTGTGCAAAATAAGGAAGCTGCTGGAG 6594
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GAATGCGAGTACTCAAGCCCGCGTGTAAATGACGGTAGTGGCAATCTTTTGTATGAA 6654
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATTGGATCTTCAAGCCAGCAAAACAGTCACTCAGCAAGACAGACTTACAAT 6714
Qy 1112 LeuProGluSerArg 1116
Db 6715 CTCATGCGATCAGAC 6729
RESULT 8
LOCUS AX459891 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 34 from Patent WO0240668.
ACCESSION AX459891
VERSION AX459891.1 GI:21725654
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Techopp,J. and Martinon,F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 34 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source 1..6939
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="NALP13/Py17.cdna"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 5724.50 Matches: 1106
Percent Similarity: 90.53% Conservative: 3
Best Local Similarity: 90.29% Mismatches: 7
Query Match: 80.92% Indels: 109
DB: 6 Gaps: 3
US-10-066-521-6 (1-1344) x AX459891 (1-6939)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyxGlyLeuGlnTrpCysLeuTrpGlu 20
Db 3055 ATGGAAGGAGACAAATCGCTCACCTTTCAGCTACGGCTGCAATGGTCTCTATGAG 3114
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLysLysLysSerSerGluSer 40
Db 3115 CTAGACAAGGAAGAAATTCAGACATTCAGAAATTAAGGAAGAAATCTTCAGAAATCG 3174
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 3175 ACCACATGCTCTATTCCACAGATTGGAATCGAAGATGCCAACGCGGAATCTGCGCCTC 3234


```
Db 5395 ATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACTGGGAGGACCCACCTGAAGAA 5454
Qy 760 GluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeu 779
Db 5455 GAGGATGTAAGATGGCGTGTGAAGCTTAAACACACCCAAATGTTGTGGAGTCTTTG 5514
Qy 780 ArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThr 799
Db 5515 AGGCTGGATGCTGTGGATGACCCATGCTTACCTGAAGATCTCCCAATCCTTACG 5574
Qy 800 ThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyVal 819
Db 5575 ACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAAGGTGACACACAGGAGTA 5634
Qy 820 MetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGlu 839
Db 5635 ATGCTCTCAGTGATGCTTGGAGTCTCCAGTGGCCCTGCAGAGCTGATACTGGAG 5694
Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
Db 5695 GACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGG 5754
Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeu 879
Db 5755 AGCTTGACACACCTGTGCTATCCCAACACAGCCTGGGGAAACGAAGGTGTAATCTACTG 5814
Qy 880 CysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHis 899
Db 5815 TGTCCATCATGAGCTTCCCCACTGTGTAGTCTGCAGAGCTGTAGTCTGAATCAGTCCAC 5874
Qy 900 LeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThr 919
Db 5875 CTGGACACGGCTGGCTGTTTCTGCACCTGGCTTATGGGTAACTCATGGCTGACG 5934
Qy 920 HisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluVal 939
Db 5935 CACCTGAGCCTTAGCATGAACCTGTGAAGACAAATGCGTGAAGCTTCTGTGGAGGTC 5994
Qy 940 MetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAla 959
Db 5995 ATGAGAGAACCACTTGTGTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTCACGCC 6054
Qy 960 AlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAsp 979
Db 6055 CGGTCTGTGAGAGTCTGCTGTGTGATCTCGAGGAGCAGACACCTCAAGAGCCTGGAT 6114
Qy 980 LeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGln 999
Db 6115 CTCACGGACAATGCCCTGGGTGACGGTGGGGTTGCTGCACCTGTGCGAGGAGCTGAAGCA 6174
Qy 1000 LysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCys 1019
Db 6175 AGAACAGTGTCTGACGAGACTCGGGTGAAGGACATGTGGACTGACTTCTGATTGCTGT 6234
Qy 1020 GluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGln 1039
Db 6235 GAGGCACTCTCTTGGCCCTTCTCGTGCACCGCATCTGACCCAGTCTAAACCTGGTGAG 6294
Qy 1040 AsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSer 1059
Db 6295 AATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGTGGCCTGTGCCACGCT 6354
Qy 1060 AsnLeuGlnIleIleGly----- 1065
Db 6355 AACTTACAGATAATTTGGCAATGACTCTGAAGAAAATGACGTTCTTCGAGAACTCTGCTCTA 6414
Qy 1065 ----- 1065
Db 6415 CTAGTTTTGCTTAAGTCACTGTTTCCAAGAACCTTATCAATGACATTAAGGAGGAACCTTA 6474
Qy 1065 ----- 1065
Db 6475 CTGTACCTCCCAAAACCCCTATAACACGACCCGGCACAGAGACTCTGTAGGAAGGAATTCAT 6534
```

```
Qy 1066 -----LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGlu 1080
Db 6535 GGATGGACTGAAAGGCTGTGAAATGGCAGTACCCCTGTGCAATAGAAAGCTGCTGGAG 6594
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 6595 GAAGTGAGCTACTCAAGCCCGAGTCGTAATTAAGACGTAGTTGGCATTCCTTTGATGAA 6654
Qy 1101 AspAspArg-----HisLysIleGlyLeuThrPheArg 1111
Db 6655 GATGACCGATTTGATCTTCAAGCCGCAAAACAGTCACCTCAGCAAGACAGACTTACAAT 6714
Qy 1112 LeuProGluSerArg 1116
Db 6715 CTCATGGCATCACAG 6729

RESULT 9
AX704823 3830 bp DNA linear PAT 04-APR-2003
LOCUS Sequence 3 from Patent EP1285964.
DEFINITION AX704823
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Weiss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
SCHERING AKTIENGESSELLSCHAFT (DE)
FEATURES
source 1. 3830
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. NO.: 0 Length: 3830
Score: 5693.50 Matches: 1087
Percent Similarity: 95.87% Conservative: 5
Best Local Similarity: 95.43% Mismatches: 12
Query Match: 80.34% Indels: 35
DB: 6 Gaps: 2

US-10-066-521-6 (1-1344) x AX704823 (1-3830)
Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGAGACAATAATCGTCACCTTTTCCAGCTACGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysLysSerSerGluSer 40
Db 61 CTACACAAGGAAGAAATTCAGACATTCAGGAATTTACTAAAGAAGAAATCTTCAGAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTTCACACAGTTTGAATCGAATCGAATCGCAACGCGGAATGCTCGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCCTTGATGAGTATTATGGAGCATCGCTGGCTTGGCTTACGTCATAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaAlaArgAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCCCTCTCGAGAGAGCGACGGGATGACATGATAAATTCACCAGAA 300
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGln--- 115
Db 301 GATCCTGAAGCAACAGTACTGACCAAGGACCAAGCAAGGAAAAAAGTGCCAGAAATAAA 360
```


QY 115 ----- 115
Db 361 TATGSCATGACTAAAGCTTATCTTGGGGGTGTCGACATCTCTGACTCGAATAATAAACAAC 420
QY 116 ----- -GlutIleSerGlnAlaMetGluGlnGluGly 125
Db 421 AAGTATGTTGAATTCATTCTTCTTTCAGAAATTTTCACAAAGTATGGAACAAGAAGGT 480
QY 126 AlaThrAlaAlaGluThrGluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSer 145
Db 481 GCCACAGCAGCAGACAGACAAGAAATGAGGATGACATGGGACTACACAGAGT 540
QY 146 HisValMetThrLysPheAlaGluGluGluAspValArgArgSerPheGluAsnThrAla 165
Db 541 CACGTGATGACCAAAATTCGCTGAGAGGAGAGATGTAGCTGCTAGTTTGAATAACACTGCT 600
QY 166 AlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPhe 185
Db 601 GCTGACTGGCCGGAAATGCAAAAGTGGCTGGTGGCTTTTGATTCAGACCAGGTGGGGCTTC 660
QY 186 ArgProArgThrValValLeuHisGlyLysSerGlyLysSerGlyLysSerAlaLeuAlaArg 205
Db 661 CGGCTTCGCACGGTGGTCTTCACCGAAAGTCAGAAATGGGAATCGGCTCTAGCCAGA 720
QY 206 ArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhe 225
Db 721 AGGATCGTGTGCTGGGCGCAAGGTGACTCTACCAAGGAAATGTTCTCTCTACGCTTTC 780
QY 226 PheLeuProValArgGluMetGlnArgLysLysGluSerSerValThrGluPheIleSer 245
Db 781 TTCCTCCCGTTAGAGAGATCGACGGGAGAGAGAGAGAGATGTCACAGAGTTCACTTCC 840
QY 246 ArgGluTrpProAspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeu 265
Db 841 AGGAGTGGCCAGACTCCAGGCTCCGGTGCAGAGATCATGTCCCGACAGCAAAAGGCTG 900
QY 266 LeuPheIleIleAspGlyPheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeu 285
Db 901 TTGTTCATCATGTAGCGGTTTCGATGACCTGGGCTCTGTCTCCAAACAATGACACAAAGCTC 960
QY 286 CysLysAspTrpAlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLys 305
Db 961 TGCAAGAGACTGGGTGAGAACAGCTCCGTTTCACTTACCTCATACGAGTGTGCTGAGGAAG 1020
QY 306 ValLeuLeuProGluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeu 325
Db 1021 GTCCTGCTCCCTGAGTCTCTCTCATCGTCAACCGTCAGAGACGTGGGCACAGAGAAGCTC 1080
QY 326 LysSerGluValValSerProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArg 345
Db 1081 AAGTCAGAGGTGCTGTCTCCCGGTTACTGTGTAGTGTAGAGAAATCTCCGGGGAACAAAGA 1140
QY 346 IleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAla 365
Db 1141 ATCCACTTGCTCCTTGAGCGCGGATTTGGTGAGCATCAGAAGACACAAGGGTTGCGTGGC 1200
QY 366 IleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuIle 385
Db 1201 ATCATGAACAACCGTGAGCTCTCGACAGTCCAGGTGCGCCGCGTGGGCTCTCTCATC 1260
QY 386 CysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThr 405
Db 1261 TGGCTGGCCCTTGACAGTGCAGAGCGTGGGGGAGAGCGTCGCCCTTCAACCAAAAGC 1320
QY 406 LeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArg 425
Db 1321 CTCACAGGCTTGACCGCGCTTTTGTGTTCATCAGCTCACCCCTCGAGCGGTGCTCCGG 1380
QY 426 ArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCysArgMetAlaValGlu 445
Db 1381 CGCTGCTCTCAATCTGGAGAAAGAGTGTCTCTGAAGCGCTTCTGCCGTATGGCTGTGGAG 1440

QY 446 GlyValTrpAsnArgLysSerValPheAspGlyAspAspLeuMetValGlnGlyLeuGly 465
Db 1441 GGAAGTGGGAATAAGGAAGTCAGTGTGTGACGGGTGACGACCTCATGGTTCAAGGACTCGGG 1500
QY 466 GluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGlu 485
Db 1501 GAGTCTGAGCTCGTGTCTGTGTTTCATGTAAACATCTCTTCCACAGACGCACTGTGAG 1560
QY 486 GluTyrTrpThrPhePheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpVal 505
Db 1561 GAGTACTACACTTCTTCCACTCAGTCTCCAGGACTTCTGTGTCGCGCTTGTACTACGTG 1620
QY 506 LeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArg 525
Db 1621 TTAGAGGGCCCTGGAATTCGAGCCAGCTCTGTGCCCTCTGTACCTTGAGAAAGACAAAGAGG 1680
QY 526 SerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeu 545
Db 1681 TCCATGGAGCTTAAACAGGACGCTTCCATATCCACTCGCTTGGATGAAGCGTTCTTG 1740
QY 546 PheGlyLeuValSerGluAspValArgProLeuGluValLeuLeuGlyCysProVal 565
Db 1741 TTTGGGCTCTGTGAGCGAAGACGTAAAGAGGCCACTGGAGGCTCTGTGGGCTGTCCGCTT 1800
QY 566 ProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsn 585
Db 1801 CCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGGTCTCTGTGTGGGTACGAGCCTTAAT 1860
QY 586 AlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLys 605
Db 1861 GCCACCACCCAGAGACACCTTGGAGCGCTTCCACTGTCTTTTCGAGACTCAAGACAAA 1920
QY 606 GluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsn 625
Db 1921 GAGTTTGTTCGCTTGGCAATTAACAGCTTCAAGAAAGTGTGGCTTCCGATTAACACAGAAC 1980
QY 626 LeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArg 645
Db 1981 CTGGACTGTGATGACATCTTCTTCTGCTCCAGCATCTCCGATTTTGGCGAAATTCGG 2040
QY 646 ValAspValLysGlyIlePheProArgAspGluSerAlaGluAlaCysProValValPro 665
Db 2041 GTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCTCTGTGGTCCCT 2100
QY 666 LeuTrpMetArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeu 685
Db 2101 CTATGGATGGGATTAAGACCTCTCATTTGAGGAGCAGTGGGAGATTTCTGTCTCATGCTT 2160
QY 686 GlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAla 705
Db 2161 GGCACCCACCCACCTTGGCGAGCTGGACCTGGGCGAGCAGCATCTTGACAGAGCGGCCC 2220
QY 706 MetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPhe 725
Db 2221 ATGAAGACCTGTGTGCCAAGCTGAGGCATCCCACTGCCAAGATACAGACCTTGATGTTT 2280
QY 726 ArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArg 745
Db 2281 AGAAATGCACAGATTACCCCTGTGTGTCAGACCTCTGGAGAAATCGTCAATGGCCAACCGT 2340
QY 746 AsnLeuArgSerLeuAsnLeuGlyThrHisLeuLysGluGluAspValArgMetAla 765
Db 2341 AACCTAAGATCCCTCAACTTGGGAGGACCCCACTGAAGAGAGGATGTGAAGATGGCG 2400
QY 766 CysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGly 785
Db 2401 TGTGAAGCCTTAAACACCCAAATGTTTGTGGAGTCTTTTGAGGCTGGATTCTGTGGA 2460
QY 786 LeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThrSerProSerLeuLys 805
Db 2461 TTGACCCATGCTGTGTACCTGAAGATCTCCAAATCTTTACGACCTCCCCCAGCCTGAAA 2520
QY 806 SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSerAspAla 825


```

Db      2521 TCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAGGAGTAATGCTCTCAGTGATGCC 2580
Qy      826 LeuArgValSerGlnCysAlaLeuGlnLysLeuLeuLeuGluAspCysGlyIleThrAla 845
Db      2581 TTGAGAGTCTCCAGTGGGCTCGAAGAGTGATCTAGGAGGACTGTGGCATCACAGCC 2640
Qy      846 ThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCys 865
Db      2641 ACGGGTTGCAGAGTCTGGCCTCAGCCCTCGTCAGCAACCGAGGCTTGACACACCTGTGC 2700
Qy      866 LeuSerAsnAsnSerLeuGlyAsnGlnGlyValAsnLeuLeuLeuCysArgSerMetArgLeu 885
Db      2701 CTATCAACAACAGCCTGGGAAACAAGGTGTAATCTACTGTGTGCATCCATGAGGCTT 2760
Qy      886 ProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCys 905
Db      2761 CCCCACTAGTCTGCAGAGGCTGATGCTGATCATGTCACCTGGACACCGCTGGCTGT 2820
Qy      906 GlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMet 925
Db      2821 GGTTCCTTGCACCTTGGCTTATGGGTAACCTCATGGCTGACGACCTGAGCCTTAGCATG 2880
Qy      926 AsnProValGluAspAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCys 945
Db      2881 AACCCCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACCATCTTGT 2940
Qy      946 HisLeuGlnAspLeuGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeu 965
Db      2941 CATCTCCAGGACCTGGAGTGTGTAAGTGTATCTCATCCCGCGCTGTGTGAGAGTCTG 3000
Qy      966 SerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeu 985
Db      3001 TCCTGTGTGATCTCGAGGAGCAGACACCTGAAGACCTGGATCTCAGCGACAATGCCCTG 3060
Qy      986 GlyAspGlyValAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThr 1005
Db      3061 GGTGACGGTGGGGTGTCTGCTACGTGCGAGGAGCTGAAGAGCAAAAGAACAGTGTCTGACG 3120
Qy      1006 ArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAla 1025
Db      3121 AGACTCGGGTGAAGGCATGTGACTGACTCTGATTCGTGTGAGGCACTCTCTCTGGCC 3180
Qy      1026 LeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAspPheSerProLys 1045
Db      3181 CTTTCTGCAACCGGCATCTGACCACTGTAACCTGGTGCAGATAACTTCAGTCCCAAA 3240
Qy      1046 GlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleGly 1065
Db      3241 GGAATGATGAAGCTGTGTTCGGCCCTTGCCTGTCACAGCTTAACCTTACAGATAATTGGG 3300
Qy      1066 LeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeu 1085
Db      3301 CTGTGGAATGGCAGTACCTGTGCAATAAGAAAGCTGCTGGAGGAAGTGCAGCTACTC 3360
Qy      1086 LysProArgValIleAspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db      3361 AAGCCCGGAGTGTAAATGACGGTAGTGGCAATCTTTTGTATGAAGATGACCGGTAC 3417

RESULT 10
AC011470/c
LOCUS      AC011470
DEFINITION Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.
ACCESSION AC011470
VERSION    AC011470.5
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 157141)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission

```

```

JOURNAL REFERENCE 2 (bases 1 to 157141)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 157141)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 15, 2000 this sequence version replaced gi:7690109.
Draft sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
    source
        1. 157141
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /clone="CTC-490M10"
ORIGIN
Alignment Scores:
Pred. No.: 3,87e-199 Length: 157141
Score: 2814.00 Matches: 658
Percent Similarity: 61.25% Conservative: 69
Best Local Similarity: 55.43% Mismatches: 175
Query Match: 39.78% Indels: 290
DB: 9 Gaps: 34
US-10-066-521-6 (1-1344) x AC011470 (1-157141)
Qy 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGlu 154
Db 42000 GGACATGGAGGTGCACATGGGACTACAGAGTCACGTGATGACCAATTCGCTGAGGAG 41941
Qy 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 41940 GAGGATGTAGCTCGTACGTTTTTGAACACACGTGCTGCTGCTGCGCGAAATGCAACGCTG 41881
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db 41880 GCTGGTCTTTTGAATTCAGACCGGTGGGGCTTCGGCTTCGACGGTGGTTCGACGGA 41821
Qy 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 214
Db 41820 AAGTCAGGAATGGGAATCGGCTCTAGCCAGAGGATCGTGTGCTGGCGCAAGGT 41761
Qy 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
Db 41760 GGACTCTACAGGGAATGTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATCGACGG 41701
Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db 41700 AAGAAGGAGAGCAGTGTTCACAGAGTTCATCTCCAGGAGGTGGCCAGACTCCACGGCTCCG 41641
Qy 255 ValThrGluIleMetSerArgProGluArgLeuPheIleLeuAspGlyPheAspAsp 274
Db 41640 GTGACGAGATCATGTCGCCAGCAAGAGGCTGTGTTCATCATGCGGTTTCGATGAC 41581
Qy 275 LeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 41580 CTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAGACGCT 41521
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 41520 CCGTTCCCTCATPACGAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTCTGATC 41461

```

```

Unpublished
REFERENCE 2 (bases 1 to 157141)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 157141)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 15, 2000 this sequence version replaced gi:7690109.
Draft sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
    Location/Qualifiers
        1. 157141
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /clone="CTC-490M10"
ORIGIN
Alignment Scores:
Pred. No.: 3,87e-199 Length: 157141
Score: 2814.00 Matches: 658
Percent Similarity: 61.25% Conservative: 69
Best Local Similarity: 55.43% Mismatches: 175
Query Match: 39.78% Indels: 290
DB: 9 Gaps: 34
US-10-066-521-6 (1-1344) x AC011470 (1-157141)
Qy 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGlu 154
Db 42000 GGACATGGAGGTGCACATGGGACTACAGAGTCACGTGATGACCAATTCGCTGAGGAG 41941
Qy 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 41940 GAGGATGTAGCTCGTACGTTTTTGAACACACGTGCTGCTGCTGCGCGAAATGCAACGCTG 41881
Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db 41880 GCTGGTCTTTTGAATTCAGACCGGTGGGGCTTCGGCTTCGACGGTGGTTCGACGGA 41821
Qy 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 214
Db 41820 AAGTCAGGAATGGGAATCGGCTCTAGCCAGAGGATCGTGTGCTGGCGCAAGGT 41761
Qy 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
Db 41760 GGACTCTACAGGGAATGTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATCGACGG 41701
Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db 41700 AAGAAGGAGAGCAGTGTTCACAGAGTTCATCTCCAGGAGGTGGCCAGACTCCACGGCTCCG 41641
Qy 255 ValThrGluIleMetSerArgProGluArgLeuPheIleLeuAspGlyPheAspAsp 274
Db 41640 GTGACGAGATCATGTCGCCAGCAAGAGGCTGTGTTCATCATGCGGTTTCGATGAC 41581
Qy 275 LeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 41580 CTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAGACGCT 41521
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 41520 CCGTTCCCTCATPACGAGTCTGCTGAGGAAGGTCTGCTCCCTGAGTCTCTCTCTGATC 41461

```

QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
DB 41460 GTACCGTCAGAGAGCTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCCCCGTAC 41401
QY 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
DB 41400 CTGTTAGTTAGAGGAATCTCGGGGAAACAAGAATCCACTTGCCTTGTAGCGGGGATT 41341
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
DB 41340 GGTGAGCATCAGAGACACAAGGGTTGGTCGGATCATGAACAACCGGTGAGCTCTCGAC 41281
QY 375 GlnCysGlnValProAlaValAlcYSerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
DB 41280 CAGTGCAGGTCGCCCGCTGGGCTCTCTCATCTGGGTGGCCCTGCAGCTGCAGAGCTG 41221
QY 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
DB 41220 GTGGGGAGAGCGTCGCCCTTCAACCAACGCTCAAGGCTGCACCGCGCTTTGTG 41161
QY 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgVal 434
DB 41160 TTTTATCAGCTACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGATT 41101
QY 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
DB 41100 GTCTGAAGCGCTTCGCCGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAGTGTTT 41041
QY 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
DB 41040 GACGCTGACGACCTCATGGTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGCTTTTAC 40981
QY 475 MetAsnIleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSer 494
DB 40980 ATGAACATCTCTTCCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGT 40921
QY 495 LeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAla 514
DB 40920 CTCCAGGACTTCTGTGCCGCTTGTACTAGTGTGTAGGGGCTTGGAAATCGAGCCAGCT 40861
QY 515 LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
DB 40860 CTCTGCCCTCTGTAGTGTAGAACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTC 40801
QY 535 HistLeHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
DB 40800 CATATCCACTCGCTTGGATGAAGCGTTCTTGTGTGGCTCGTGAGCGAAGACGTAAGG 40741
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
DB 40740 AGGCCACTGGAGGTCCTGTGGGCTGTCCGTTCCCTGGGGGTGAAGCAGAGCTTCTG 40681
QY 575 HistTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
DB 40680 CACTGGGCTCTCTGTGGGTGAGCAGCGCTTAATGCCACCACCCCGAGGAGACCCCTGGAC 40621
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
DB 40620 GCCTTCCACTGCTCTTTTCGAGACTCAAGACAAAGAGTTGTGTGCTGGCATTAACACAGC 40561
QY 615 PheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCys 634
DB 40560 TTCCAGAGAGTGGCTTCCGATTAACAGNAACCTGGACTGTATAGCATTTTCTCTCTGC 40501
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
DB 40500 CTCCAGCACGTCTCGGTATTTGGGAAATTCGGGTGGATGCAAGGAGTCTTCCCAAGA 40441
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667
DB 40440 GATGAGTCCCGTGAGGCATGCTCTGTGGTCCCTCTATGGTGAGTACCCCGAGCAGTTTGA 40381
QY 667 ----- 667

DB 40380 TCCTATGCCGTGTGCTGAGCTCTGTGCTCTACTACTGCTGGGACTTGACATGACTTCCAGGA 40321
QY 668 ---MetArgAspLysThrLeuIleGlu-----GluGlnTrp--- 678
DB 40320 ACTTCAAGGTCCAGAGAAATTTCTTCAGGATGAATGGCCAGATGAGCTCCAGCTGGCTA 40261
QY 679 -----GluAsp-----PheCysSerMetLeuGly-----Thr 687
DB 40260 AAATGCCAGGACCGGAATCTGGCTCATTTGTCACACTGGGAAAGGCTTTTCAATACA 40201
QY 688 HisProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArg----- 704
DB 40200 GGGCCCGCCCTAAGA-----CTAAAGGAAGGCTCTGGGT 40165
QY 705 -----AlaMet 706
DB 40164 TGAACCCAGGAATATGTGGTTTATTTTATTTTAAAGTCAGAGTCTCACTCTGTGTC 40105
QY 707 LysThrLeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArg 726
DB 40104 AGCTGGAGGTAGTGGCAGCATCTCGGCTCACTGCAAACTCCACCTCCAGGCTCAAGA 40045
QY 727 AsnAlaGlnIle-----ThrProGlyValGlnHis----- 736
DB 40044 GATTCCTCTCCCTCAGGCTGCCGAGGAGCTGGGACTTACAGGCAATGACACCATGCCAG 39985
QY 737 LeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHis 756
DB 39984 CTAGTTTTGTAAATTTTAGTAGAGATGGGATTTTCGC-----CACGTTGGCCAGGCTGGT 39931
QY 757 LeuLysGluGluAspValArg----- 763
DB 39930 CTCGAAT-CTTGACCTCAGGTGATCCACCGCTCAGCCTCCCAAAGTCTGGGATTAC 39872
QY 764 -----MetAlaCysGluAlaLeu----- 769
DB 39871 AGGCATCAGCAAGCCCTCCCATGATGTGACCGCTCAGGTTGTGGACTGTATTATGG 39812
QY 770 -----LysHisProLysCys 774
DB 39811 TATGTGTCCAGGGCGGTGTGTCCAGTAGAGGTTCTGTGATGGAATATTCATGTGC 39752
QY 775 Leu-----LeuGluSerLeuArgLeu----- 781
DB 39751 CTGTCCAGTAGACAGCCACGACGCTCATAGCCTGTTGAGCATCTGAACATGTGTAG 39692
QY 782 ---AspCysCysGlyLeuThrHis-----AlaCysTyrLeu 792
DB 39691 TGAGACTGTGGAGTGAAACTCATCTCACTCCATTTCAACTCATTTGTATTTT-TTCTT 39633
QY 793 LysIleSerGlnIleLeuThrThrSerProSerLeuLys----- 805
DB 39632 TTTTTCAGACAGAGTCTCACTGTGTCCCAAGCTGGAGTACAGCAGCATGATCTTTGCT 39573
QY 806 -----SerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMetProLeuSer 823
DB 39572 CACTGCAATCTCTGCCCTCCGGGTTC-----AGCGATTCTCCTGA-CTCAGCCTCTCA 39520
QY 824 AspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle-----LeuGluAsp 840
DB 39519 AGTAGCTGGGATATAGGAGTGGCCACCATCTGCTGCTAAATTTTGTATTTTAGTAGAG 39460
QY 841 Cys-GlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg-- 859
DB 39459 ACAGGGTTTCAAC-----ATGTTGGCCAGGCTGATCTCGAACTCCTG 39418
QY 860 -----SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLe 878
DB 39417 ACCTCAGGTGAGCCACTGCC-TTGGCTCTCCAAAGTGTGGGATTTATAGGCATAGCTG 39359
QY 878 uLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArg----- 892

QY 295 ProPheThrLeuIleAArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
DB CCGTTACCCCTCATACGCGAGTCTGCTGAGGAGGTCTCTGCTCCCTGAGTCTCTCTGATC 22797
QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
DB GTCAACCGTCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTGCTGTCTCTCCCGTTAC 22857
QY 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
DB CTGTGTAGTTAGAGGAATCTCCGGGGAAACAAAGATCCACTTGTCTCTTGAGCGCGGAT 22917
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
DB GGTGAGCATCAGAGACACAGAGGTGGTGGTCTCATGATCAATCAACACCGTGAGTCTCTGCAC 22977
QY 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
DB CAGTGCCAGGTGCCCGCGTGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGACGTG 23037
QY 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
DB GTGGGGGAGACGTGCGCCCTTCAACCAACGCTCACAGCCCTGCAGCGCTTTTGTG 23097
QY 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgVal 434
DB TTTTATAGCTCACCCCTCAGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGATT 23157
QY 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
DB GTCTGAAGCGCTTCTGCGGTATGGCTGTGGAGGAGTGTGGATAGGAAGTCAGTCTTT 23217
QY 455 AspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
DB GACGGTCACGACCTCATGGTTCAAGGACTCGGGAGGCTGAGCTCCGCTGCTCTTTTAC 23277
QY 475 MetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPheHisLeuSer 494
DB ATGACATCTCTTCCACAGACGACCTGTGAGAGTACTACACTTCTTCCACTCAGT 23337
QY 495 LeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAla 514
DB CTCCAGGACTTCTGTGCGCCTTGACTACTGTTAGAGGCGCTGGAAATCGAGCCAGCT 23397
QY 515 LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
DB CTCTGCCCTCTGTACGTGTGAAGACAAAGAGGTCCATGAGGCTTAAACAGGCGGCTTC 23457
QY 535 HisIleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
DB CATATCCACTCGCTTTGTGTAAGCGTTTCTGTGTGGCTCTGTGAGCGAAGACGTAAG 23517
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
DB AGGCCACTGGAGGTCTGCTGGGCTGCTCCGCTTCCCTCGGGGTGAAGAGCTTCTG 23577
QY 575 HisTrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
DB CACTGGGTCTCTCTGTGGGTGACAGCCCTAAATGCCACCAACCCCGAGAGACACCTCGAC 23637
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
DB GCCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTGCTTGGCATTAACAGC 23697
QY 615 PheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCys 634
DB TTCCAAGAAGTGTGGCTCCGATTAAACAGAACCTGGACTGTAGCATCTCTCTCTCTG 23757
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
DB CTCAGACACTGTCTCGTATTTTGGGAAATTTGGGTGGATGTCAAGAGGATCTTCCCAAGA 23817
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667

DB 23818 GATGAGTCCGCTCAGGCAATGCTCTGTGTCCTCTATGG 23856
RESULT 12
AY196362 3470 bp mRNA linear ROD 18-DEC-2003
LOCUS Mus musculus strain CBA/J WATER protein (Water) mRNA, complete cds.
DEFINITION AY196362
ACCESSION AY196362.1 GI:37781353
VERSION 1 (bases 1 to 3470)
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3470)
AUTHORS Cheng,H., Gao,J., Zhang,Y., Huang,Z., Liu,H., Teuscher,C. and Ma,R.Z.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2002) Human and Animal Genetics Division,
Institute of Genetics & Developmental Biology, Chinese Academy of
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China
FEATURES
Location/Qualifiers
1..3470
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CBA/J"
/db_xref="taxon:10090"
/chromosome="7"
1..3470
/gene="Water"
9..3344
/gene="Mater"
/codon_start=1
/product="WATER protein"
/protein_id="AA052698.1"
/db_xref="GI:37781354"
/translation="MGPPKDSKAILKARLEEEKSERKMTSPENDSKSIOKDQPE
QQTSESTMGPPKDSKAILKARLEEEKSESTSPENSKAILKDSGSEVQAS
ERKMTSPENDSKSIOKDQPEQSETLQSKBEDEVTEADKNGDLDQYKAHVQAK
PSTVDLHYDPSKMLSDAFQYKTFQPTIILHGRFVGKSAARSIVLWAGGK
LFQKSFVIFSVRIKWTSESLAQIAKECPDWDVPTKMSQPERLLFVLDGDD
MDSLQHDMTLSRDWKDQPIYILMYSLRKALLPQSLITTRNTGLEKLKSMVVS
PIYLVEGLSASRRSOLVLENTSNESDRIOVFHSLIENHOLFQDQAPSVCSVCEAL
YQKLKGRCTLPCTLTGLYATLPHVQTLKRPQSNALSQEQITLVGCMMAEGV
WTMSVYFDDDLKNSYKESLEALFPHNILLQVHNSQCYVFSHLSQDFFAALY
VLEGLEBWNQHFCTENQSRIMEVKTDDTLRLGMRKFLGLMKNKDLTKTLVLEYP
VPTVEQLQHWVSLIAQQVNGTSPMDTLDAFYCLFESQDEEFVGALKRFOVLLI
NOKMDLVSSYCLKHCONLKAIRVDIRLLSDNTLELCPVTVTOCTCKPILMBWG
NFCSVLSGLRLKEULDIGDSILSQAMKILCLELRNQCRIQKLTFSKAEVYSGLKH
WKLPSNQMLKYNLGNTPMKDDDMKLAACEALKHPKCSVETLRDSCELTIIGYEMIS
TLISITTRLUKLSLAKNRVGVKSMISLGNALSSMLQKILNCGNLTIPASCHLLVS
ALFSNQNLTHLCLSNNSLGTGVQQLCOFLRNPECALQRLILNHCNIVDDAYGFLAMR
LANNTKLTHLSLTMPNPGDMKLCEALKEPTCYLOELVDCOLTQNCEDLACMI
TTTTKLSLDIGNNALGDKGVITLCEGLKQNNSSLRRLGLGACKLTSNCCALSLAIS
CNPHLSNLVLKNDSTSGMLKLSAFQCPVSNLGTIGLWKQVEYVARVROLEVEFV
KPHVVIDGWYASDEDDRNWNKN"

Alignment Scores:
Pred. No.: 1,36e-192 Length: 3470
Score: 2695.00 Matches: 555
Percent Similarity: 66.49% Conservative: 185
Best Local Similarity: 49.87% Mismatches: 293
Query Match: 38.10% Indels: 81
DB: 10 Gaps: 11
US-10-066-521-6 (1-1344) x AY196362 (1-3470)

QY 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPheLysGluLeuLys 34
DB 163 CAATGGTCTCTCCA-GAAAAAGACAGTAAAGCAATCTTGAAGACGCTGGATTGGAAGAG 221
QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54

Db 222 GACAGAAAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAsp 94
Db 243 ---ATGTCTCTTCAGAAAATGTCTAGTAAAGCCATCTCTGAAAAGACAGTGGATCAGAGAA 299
Qy 95 MetLysLysIleSerGln----- 100
Db 300 GTGGAAACAGGCATCAGAAAGAAAATGACTTCTCCAGAAAACGACAGTAAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGln 115
Db 360 AAAGACCAAGCAGCAGCAGGAGCAGACATCAGAAACCTTACAACTCTAAGGAAGAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 420 GAAAGTACAGAGGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
Db 444 AATGAGGTGACTTCAAGACTTACAAAGCCCATGTGATTGCTAAGTTCCGACCAAGTGTG 503
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCAGAGATGAATATTGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATGCTTTTAAACCATACAGAAAACCTTCCAGCCTCACCACTATCTACATGGAAGA 602
Qy 196 SerGlyIleGlySerAlaLeuAlaArgIleValLeuCysTrpAlaGlnGlyGly 215
Db 603 CCAGGAGTGGGAAGTACAGCTTGGCCAGAGATATTGTTCTGGCTGGGCACAGGGTAAA 662
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 663 CTCCTCCAAAAGT---TCCTTTGTCTCTCTCTCTCTGTTAGAGAAATAAAGTGACA 719
Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 720 GAGAAGAGCAGTTTGGCACAGCTGATTGCTAAGGAGTGTCCAGACTCTCTGGGATCCAGTG 779
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspLeu 275
Db 780 ACAAGATCATGTCCCAACCAAGAAAGACTCTTGTGTGTCATAGATGGCTTGGATGATATG 839
Qy 276 GlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 840 GACTCTGCTCCCAACATGATATGACATATCCAGAGACTGGAAGATGAACAGGCC 899
Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 900 ATATACATCTCTGATCAGACCTCTCTGAGGAAGGCTCTCTACCTCAGTCTCTTCTCATC 959
Qy 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
Db 960 ATTACCACCAACACAGGCTTATGAAAACCTCAAGTCAATGGTTGTTTCCCTCTCTAT 1019
Qy 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 1020 ATACTGGTTGAGGACTCTCTGCATCAGGAGATCTCAGCTGGCTCTCGAGAAACATCTCC 1079
Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
Db 1080 AATGAGTCTGATAGATACAAGTCTTCCATTCTCTGATAGAAAATCACCAGCTGTTTGAC 1139
Qy 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
Db -----

Db 1140 CAATGCCAGGCCCTCTGTGTGCTCCCTGGTCTGTGAGGCTCTGCAGCTACAGAGAAA 1199
Qy 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
Db 1200 CTGGAAAGAGATGCACCCCTACCTCCGACAGACTCTCACCGGTTGTATGCCACGTTGGTG 1259
Qy 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgVal 434
Db 1260 TTTTCCACAGCTCACCTTGAAGAGCCCTTCCAGAGCGCTCTCAGTCAGGAAGAACAGATT 1319
Qy 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
Db 1320 ACTTACTGGTGTGTCATGATGGCAGCTGAAGAGTGTGGACCATGAGGTCCGGTGTTC 1379
Qy 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
Db 1380 TATGATGATGATCTGAAGAACTATAGCTTAAAGAGTCTGAGATCTTGGCCCTCTTTCAC 1439
Qy 475 MetAsnIleLeuLeuProAspSerHis---CysGluGluTyrTyrThrPhePheHisLeu 493
Db 1440 ATGAACATCTCTTCCAGGTTGGCCACACAGTGCAGTGTATGTTTCTCCACCTC 1499
Qy 494 SerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu---GluIleGlu 512
Db 1500 AGCCTGCAGGATTTCTTGTGCTTATATATGTTTTAGAAAGGCTGGAGGAATGGAAT 1559
Qy 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532
Db 1560 CAGCATTTTTCG-----TTCATTGAAAACCAAGGAGCATCTGGAGGTGAAGAACT 1613
Qy 533 GlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAsp 552
Db 1614 GAC---GACACTCGCCTCTTGGGATGAAGCGTTTCTTATTGTCCTCATCAACAAGAT 1670
Qy 553 ValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLys 572
Db 1671 ATCTTGAAAGACTCTGGAGGTTCTGTTTGAATATCCCGTGTATCCAACTGTTGAGCAGAAG 1730
Qy 573 LeuLeuHisTrpValSerLeuLeuGlnGlnProAsnAlaThrThrProGlyAspThr 592
Db 1731 CTCCAACACTGGGTCTCTCTGATAGCTCAGCAGGTCAATGGCACCACCCCATGGACACC 1790
Qy 593 LeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeu 612
Db 1791 CTGATGCTCTTCTATTGCTATTGAGTCTCAGGATCAGAGATTTGTTGGCGGGCTCTC 1850
Qy 613 AsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSer 632
Db 1851 AAACGCTTCCAAAGAGTGTGGCTGCTGATTAACCAAGAGATGGACTTGAAGGTCTCTTCC 1910
Qy 633 PheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePhe 652
Db 1911 TACTGTCTCAAGCACTGTGAACTTGAGGCAATCCGGTGGATATCAGAGACCTCTCTC 1970
Qy 653 ProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---MetArgAspLys 671
Db 1971 TCGGTAGATAATACTCTCGAGCTGTGCCCTGTGTTACTCTCCAGGAGACACAATGTAAAG 2030
Qy 672 ThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeu 691
Db 2031 CCCTCTCTCATGAGTGTGGGAAACTTCTGTGTGTGTGGCAGCCTCCGGAACCTG 2090
Qy 692 ArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAla 711
Db 2091 AAGAGCTGAGACTTGGCGGACAGCATCTCAGTCAACGGCCCATGAAGATCTGTGCTC 2150
Qy 712 LysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThr 731
Db 2151 GAGCTGGGAATCAGTCTCTCAGAAATACAGAACTGACGTTTAAAGAGTCAGAGGTAGTG 2210
Qy 732 ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsn 751
Db 2211 TCTGGCTGAAACATCTCTCGAAAGCTCTCTTTTAGCAATCAAAACTTAAAGTACCTCAAT 2270

```
Qy 752 LeuGlyGlyThrHisLeuLeuGluaspValargMetAlaCysGluAlaLeuLeuHis 771
Db 2271 CTAGGAAACATCCCATGAAGATGATGACATGAAGTTAGCTCGAAGCGCTGAACAT 2330
Qy 772 ProLysCysLeuLeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyr 791
Db 2331 CCAAGTGCTCGTGAGACTCTGAGTTGATTCCTGTGAGTTAACCATCATTTGGCTAT 2390
Qy 792 LeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGly 811
Db 2391 GAGATGATCTCCAGCTCTTATTATCAACACCAAGCTGCTGAGCTGAGCTGAGCTG 2450
Qy 812 AsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCys 831
Db 2451 AATAGAGTGGAGTAAAGCATGATATCCCTGGGAATGCTTGAAGTAGCTCAATGTT 2510
Qy 832 AlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeu 851
Db 2511 CTACTGCAAAAGTTGACTGAGCAACTGTGGCTCTACACTGCCAGCTGCCACCTCTG 2570
Qy 852 AlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeu 871
Db 2571 GTCTCAGCCCTTTTCAGCAACACCAAGCTTGACACACCTGTGCTGTCAACACAGCCTG 2630
Qy 872 GlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGln 891
Db 2631 GGGACTGAAGAGTGCAACAGCTGTGCTGCTGCTGAGGATTCAGAACTGCTCTCCAG 2690
Qy 892 ArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAla 911
Db 2691 CGGCTGATCACTCACTGCAACATTTAGATGATGCTTATGCTTCTTGGCAATGAGA 2750
Qy 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsn 931
Db 2751 CTTGCAAAACACAAAGCTGACCCAGCTGAGCTGACCTGACCTGACCTGACCTGAGG 2810
Qy 932 GlyValLysLeuLeuGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGCTACTGTGTGAAGCTTTTAAAGAACTTACTTGTATACCTTCAAGAACTG 2870
Qy 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTAGTGAGCTGCCAACTCACACAACTGCTGCGAGGACCTGGCTGTATGATCACACA 2930
Qy 972 SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAAGCAATTTAAAGTTTGGATCTTGGTAAACAACGCCCTGGGTGACAAAGGAGT 2990
Qy 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuAla 1011
Db 2991 ACCCTGTGTAGGAGCTGAGCAAAATAACAGCTCCCTGAGGAGACTGGGTGGGGGCA 3050
Qy 1012 CysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAGTTGACTTCCAATGCTGTGAGGCAATGTGATGGCCATCTCTTGCACCCCTCAC 3110
Qy 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db 3111 CTGAACAGCTTAAACCTGGTGAAGAAATGACTTCAAGTACATCGGGGATGTTGAAGCTG 3170
Qy 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyr 1071
Db 3171 TCTCGTTCANAGCTCTGCTCTTAACCTGGGATAATGCGCTGTGGAAGCAGGAGTAC 3230
Qy 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIle 1091
Db 3231 TATGCCGAGTGAGACAGACAGCTGAGGAGTTGAGTTTGTCAAGCCCAAGCTGATG 3290
Qy 1092 AspGlySerTrpHisSerPheAspGluAspAspArgHis 1104
Db 3291 GATGGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
```

```
RESULT 13
AY329487
LOCUS
DEFINITION
Mus musculus strain CBA/J MATEr protein isoform-E (Mater) mRNA,
complete cds; alternatively spliced.
ACCESSION
AY329487
VERSION
AY329487.1 GI:37595472
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 3405)
AUTHORS
Cheng,H., Zhang,Y., Huang,Z., and Ma,R.Z.
TITLE
Identification and characterization of alternative splice variants
for murine Mater
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 3405)
AUTHORS
Cheng,H., Huang,Z., Zhang,Y., Liu,H., Teuscher,C. and Ma,R.Z.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUN-2003) Human and Animal Genetics Division,
Institute of Genetics and Developmental Biology, Chinese Academy of
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China
FEATURES
Location/Qualifiers
1..3405
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CBA/J"
/db_xref="taxon:10090"
1..3405
/genes="Mater"
9..3296
/genes="Mater"
/notes="alternatively spliced"
/codon_start=1
/product="MATER protein isoform-E"
/protein_id="AA094610.1"
/db_xref="GI:37595473"
/translation="MGPPKDSKAILKARGLEEQKSERMTSPENDSKSIQKQGP
EQSTSTWGPPEKSKAILKARGLEEQKSERMTSPENSKAILKSGSEVEEQAS
ERMTSPENDSKSIQKQGPPEQSTSDNGDLDQYKAHVIAKFTDSDVLDHYSPEMKL
LSDAFPKYOKTQPHITILHGRPGVGKSALARSIVLGAQGLKQKSFVIFSVREI
KTEKSLAQILAKCEPCDSWDPVTIKMSOPERLLFVIDGLDDMDSVLQHDHMTLSRDW
KQOPIYILMYSILKALLPQSPLIITENTGLEKLSMVVSPLYILVGLSASRBSQ
LVLENSNEDRIQVHSLIENHQLFDQCAPSVCSLVCEALQKQKRCRTLPQCT
LTGLYATLVFHLTKRPSQALSQEEITLVLCMAAEGVMTMSVFDVDDKNYS
LKSEITALLFHNILLOVGNSEQCQYVFSHLSDQFFAALYVLEGLSEMNQHFCTIE
NORSIMEVKRTDTRLLGMKRFGLMKNKILKLEVLFPVPTVEQKLQHWVLI
AQQVNTSPMDTDAFYCLPESQDEFVGGALKRFOEWLLINOKMDLVKSSYCLKHC
ONLKATRVDIRLLSVNDTLEICPVVTVOETCKPLLMEMWGNFCSVLRLNELD
LGDSTLSQAMKILCLELRNQSCRIOKLTFKSAEVVSGLKLKLLFSQNLKYLNLG
NTPMKDDDKLACEALHPKSVETURLDSCELTIIGYEMISTLLISTTLRLKCLSLAK
NRVGVKMSISLGNLSSCLLQKGLDNCGLTPASPCHLLVSALFQNLTHLCLSN
SLGTEGVQQLCQFLRNPECALQRLILNHCNIVDDAYGFLAMRLANNKTLTHSLTNP
VGDGAMKLEALKEPTCYLOLELVDCOLTONCCEDLACMIITTKQLKSLDGNAL
GDKVITLCEGLKONNSLRRLGLGACKLTSCCEALSIAICNPHLSNLVKNDFS
TSGMLKLCSAFQCPVSNLGLIGLKQEYIYARVRQLEVEFVKPHVVIDGDWVASDED
DRNWVKN"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,23e-192 Length: 3405
Score: 2692.00 Matches: 557
Percent Similarity: 67.52% Conservative: 183
Best Local Similarity: 50.82% Mismatches: 294
Query Match: 38.05% Indels: 63
DB: 10 Gaps: 13
```

US-10-066-521-6 (1-1344) x AY329487 (1-3405)

```
Qy 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhenylsGluLeuLys 34
Db 163 CAATGGTCTCTCCA-GAAAAAGACAGTAAGCAATCTTGAAAGCAGCTGGATTGGAAG 221
```

QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
| | | | |
DB 222 GAACAGAAAGTCAGAAAGCACA----- 242

QY 55 ValGluCysLeuAlaLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
| | | | |
DB 242 ----- 242

QY 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
| | | | |
DB 243 ---ATGTCCTCTCAGAAAATGTCAGTAAAGCCATCTCGAAGACAGCTGGATCAGAAGAA 299

QY 95 MetLysLysIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlu 114
| | | | |
DB 300 -----GTGGAACAGGCATCAGAAAGAAA---ATGACTCTCTCAGAAAACACACAGT 347

QY 115 GlnGluIle-----SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGlu 132
| | | | |
DB 348 AAATCAATCCAGAAAGACCAAGACCAGAGCAGGAG-----CAG 386

QY 133 GluGlnGlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAla 152
| | | | |
DB 387 ACATCAGATATAGAGGTGACTTACAGACTACAGGCCCATGTGATGCTTAAGTTGCAC 446

QY 153 GluGluGluAspValArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGln 172
| | | | |
DB 447 ACAAGTGTGGATCTACAC-----TATGACAGCCACAGAGATGAAA 485

QY 173 ThrLeuAlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeu 192
| | | | |
DB 486 TTATGTGTCGTGCTTTTAAACCATACACAGAAAACCTTCCAGCCCTCACACCATATCTTA 545

QY 193 HisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgAlaGluValLeuCysTrpAla 212
| | | | |
DB 546 CATGGAAGACAGAGAGTTGGGAAGTACGCTTTGGCCAGAAATGTTCTTGGCTGGGCA 605

QY 213 GlnGlyGlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMet 232
| | | | |
DB 606 CAGGTAACACTCTTCCAAAAGATG---TCCTTGTGTCATCTTCTCTGTTAGAGAAATA 662

QY 233 GlnArgLysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGln 252
| | | | |
DB 663 AAGTGGACAGAGAGAGCAGTTTGGCACAGCTGATGTCAGAGAGTGTCCAGACTCTCTGG 722

QY 253 AlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPhe 272
| | | | |
DB 723 GATCCAGTGCAGAAAGATCATGTCCACAGAAAGACTCTTGTGTTGTCATAGATGGCTTG 782

QY 273 AspAspLeuGlySerValLeu---AsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
| | | | |
DB 783 GATGATATGGACTCTGTCTCTCAACATGATGATGACACTATCCAGAGACTGGAAGAT 842

QY 292 LysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
| | | | |
DB 843 GAACAGCCCATATACATCTCTGATGATACAGCTCTCTGAGGAAGGCTCTCTACCTCAGTCC 902

QY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
| | | | |
DB 903 TTTCTCATATTACACAGAAAACACAGGCTTAGAAAACCTCAAGTCAATGGTGTGTTTCC 962

QY 332 ProArgTyrLeuLeuValArgIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351
| | | | |
DB 963 CCCCTCTATATAGTGTGTAAGGACTGTCTGCATCAAGGAGATCTCAGCTGGTCTCGAG 1022

QY 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
| | | | |
DB 1023 AACATCTCCAATGAGTCTGTAGATAACAACTCTTCCATTCTCTGTAGAAAATCACCAG 1082

QY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391
| | | | |
DB 1083 CTGTTTGACCAATGCCAGGCCCTCTGTGTGTGCTCCCTGTGTGCTGTGAGGCTCTGCAGCTA 1142

QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
| | | | |
DB 1143 CAGAAGAAAGTCGGAAGAGATGCACCTACCCTGCCAGACTCTCAACCGGTTGTATGCC 1202

QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
| | | | |
DB 1203 ACGTTGGTGTTCACACAGCTCACCTTGAAAAGGCTTCCAGAGCGCTCTCAGTCAGGAA 1262

QY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
| | | | |
DB 1263 GAACAGATTACTCTAGTGGGTTTGTGATGATGGCAGCTGAGGAGTGTGGACCATGAGG 1322

QY 452 SerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
| | | | |
DB 1323 TCGGTGTTCTATGATGATGATCTGAAGAACTATATGCTTAAAGGAGTCTGAGATCTTGCC 1382

QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHis---CysGluGluTyrTyrThrPhe 490
| | | | |
DB 1383 CTCTTTTCCATGATGAACATCTCTTCCAGGTTGGCCACACAGTGAAGCAGTGTATGTTTC 1442

QY 491 PheHisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu--- 509
| | | | |
DB 1443 TCCCACTCAGCTGCGAGGATTTCTTGTGCTCTTATATATGTTTAGAAGGGCTGGAG 1502

QY 510 GluIleGluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeu 529
| | | | |
DB 1503 GAATGGAATCAGCATTTTTCG-----TTCATTGAAAACCAAGGAGCATCATGGAGGTG 1556

QY 530 LysGlnAlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuVal 549
| | | | |
DB 1557 AAGAGAACTGAC---GACACTCGCCTCTTGGGATGAAGCGTTTCTTATTGTCCTCATG 1613

QY 550 SerGluAspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyVal 569
| | | | |
DB 1614 AACAGATATCTTGAAGACTCTGGAGTTCTGTGTAATATCCCGTGATTCCAACTGTT 1673

QY 570 LysGlnLysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrPro 589
| | | | |
DB 1674 GAGCAGAAAGCTCCAAACACTGGGTCTCTCTGATAGCTCAGCAGGTCAATGGCACCAGCCA 1733

QY 590 GlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArg 609
| | | | |
DB 1734 ATGCACACCTCGATGCTTCTATTGTATTGTAGTCTCAGGATGAAGAGTTGTTGGC 1793

QY 610 LeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnLeuLeuAspLeuIle 629
| | | | |
DB 1794 GGGGCTCTCAACCGCTTCCAAAGAGTGTGCTCTGATTAACCAAGAGATGGACTTGAAG 1853

QY 630 AlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLys 649
| | | | |
DB 1854 GTCTCTTCTTACTGTCTCAAGCACTGTCAAGAACTTGAAGGCAATCCGGGTGGATATCAGA 1913

QY 650 GlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---Met 668
| | | | |
DB 1914 GACCTCTCTCGGTAGATAATACTCTCGAGCTGTGCCCTCTGTGTACTGTCTGAGGAGACA 1973

QY 669 ArgAspLysThrLeuIleGluGluGlnTrpGluAspPheCysSerMetLeuGlyThrHis 688
| | | | |
DB 1974 CAATGTAAGCCCTCTCTCATGGAGTGGTGGGAAACTTCTGCTGTGCTTGGCGGCTC 2033

QY 689 ProHisLeuArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThr 708
| | | | |
DB 2034 CGGAACCTTGAAGAGCTGGACTTGGCGGACAGCATCTGTAGTCAACGGGCGCATGAAGATA 2093

QY 709 LeuCysAlaLysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAla 728
| | | | |
DB 2094 CTGTGCTCGAGCTCGGAATCAGTCTGCAGAAATACAGAAGCTGACGTTTAAGAGTGCA 2153

QY 729 GlnIleThrProGlyValGlnHisLeuTrpArgIleValMetAlaLeuAsnArgAsnLeuArg 748
| | | | |
DB 2154 GAGGTAGTGTGTGGCTGGAACATCTCTGGAAGCTCTTTTATAGCAATCAAAACTTAAAG 2213

QY 749 SerLeuAsnLeuGlyGlyThrHisLeuLysGluAspValArgMetAlaCysGluAla 768
| | | | |


```
Db 2214 TACCTCAATCTAGGGAACACCTCCCATGAAGATGATGACATGAAGTTAGCTCGGAAGCG 2273
Qy Leu1ysHisPro1ysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHis 788
Db 2274 CTGAACATCCAAAGTCTCGGAGACTGAGGTTGGATTCCTGTGAGTTAACCATC 2333
Qy 789 AlaCysTyrLeu1ysHisGln1leuThrSerProSerLeu1ysSerLeuSer 808
Db 2334 ATTGGCTATGATGATCTCCAGCTTCTTATTTCAACCACAGCTTAAAGTGTCTCAGC 2393
Qy LeuAlaGlyAsn1ysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgVal 828
Db 2394 CTGGCCAAATAAGAGTGGGAGTAAAGCATGATATCCCTTTGGGAATCCCTTGAGTAGC 2453
Qy 829 SerGlnCysAlaLeuGln1ysLeu1leuGluAspCysGly1leuThrAlaThrGlyCys 848
Db 2454 TCAATGTGTCTACTGCCAAAGTTGATACTGGACAACCTGTGGCCCTCACACCTGCCAGCTGC 2513
Qy 849 GlnSerLeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsn 868
Db 2514 CACCTTCTGGTCTCAGCCCTTTTCCAGCACCAGAACTTGACACACCTGTGCTGTCAAC 2573
Qy 869 AsnSerLeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCys 888
Db 2574 AACAGCCTGGGACTGAAGGAGTGCACAGCTGTGTGAGTCTCCAGGAATCCAGATGT 2633
Qy 889 SerLeuGlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeu 908
Db 2634 GCTCTCCAGCGGCTGATCTGAATCACTGCAACATTTGTAGATGATGCTTATGGCTTCCTG 2693
Qy 909 AlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProVal 928
Db 2694 GCAATGAGACTTGCACAAACACAAAGCTGACCCACCTGAGCCCTGACCAATGAAACCCCGTA 2753
Qy 929 GluAspAsnGlyVal1ysLeuCysGluValMetArgGluProSerCysHisLeuGln 948
Db 2754 GGGGATGTGTGCAATGAGTACTGTGTGAGCTTTTAAAGGAACTACTTGTACTTCAA 2813
Qy 949 AspLeuGluLeuVal1ysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysVal 968
Db 2814 GAATGGAACTAGTGGAGTGGCAACTCACACAGAACTGCTCGAGGACCTGGCCCTGTATG 2873
Qy 969 IleSerArgSerArgHisLeu1ysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGly 988
Db 2874 ATCAACACCAAGCAATTAAGAATTGGATCTTGTACAAACGCTGCGGGTGACAAA 2933
Qy 989 GlyValAlaAlaLeuCysGluGlyLeu1ysGln1ysAsnSerValLeuThrArgLeuGly 1008
Db 2934 GGAGTCATAACCTGTGTGAGGAGCTGAAGCAAAATAACAGCTCCCTGAGGAGACTTGGG 2993
Qy 1009 Leu1ysAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCys 1028
Db 2994 TTGGGGGATTAAGTGTGACTTCCAATTTGCTGTGAGGCAATTTGTCATTTGGCCATCTTTCG 3053
Qy 1029 AsnArgHisLeuThrSerLeuAsnValGlnAsnAsnPheSerPro1ysGlyMetMet 1048
Db 3054 AACCTTACCTGCAACAGCCTTAACCTGGTGAAGAAATGACTTCAGTACATCGGGGATGTG 3113
Qy 1049 LysLeuCysSerAlaPheAlaCysProThrSerAsnLeuGln1leuGlyLeuTrp1ys 1068
Db 3114 AAGCTGTGCTCTGGTTCATGCTCTCTTAACCTGGGGATAATTTGGCTGTGGAAG 3173
Qy 1069 TrpGlnTyrProValGln1leuArg1ysLeuLeuGluValGlnLeuLeu1ysProArg 1088
Db 3174 CAGGAGTACTATGCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTTGTGCAAGCCCCAC 3233
Qy 1089 Val1leuLeuAspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3234 GTGGTGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 3281
```

RESULT 14
AX427592

```
LOCUS AX427592 3447 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from Patent WO0232955.
ACCESSION AX427592
VERSION AX427592.1 GI:21537713
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 5 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source 1. 3447
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 4,54e-192 Length: 3447
Score: 2688.00 Matches: 554
Percent Similarity: 66.49% Conservative: 186
Best Local Similarity: 49.78% Mismatches: 293
Query Match: 38.00% Indels: 81
DB: Gaps: 11
US-10-066-521-6 (1-1344) x AX427592 (1-3447)
Qy 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhelysGluLeuLeuLys 34
Db 163 CAATGGTCTCCCA-GAAAAAGACAGTAAAGCAATCTTGAAGACACCTGGATGGAGAG 221
Qy 35 LysLysSerSerGluSerThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54
Db 222 GAACAGAGTCAGAAAGCACA----- 242
Qy 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74
Db 242 ----- 242
Qy 75 SerIleSerIlePheGluAsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspAsp 94
Db 243 ---ATGTCCTCTTCAAAAATGTCTAGAGCAATCTCTGAAGACAGATGGATCAGAAGAA 299
Qy 95 MetLysLysIleSerGln----- 100
Db 300 GTGGACAGCGCTCAGAAAGAAAAATGACTTCTCCAGAAAAACGACAGTAATCAATCCAG 359
Qy 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115
Db 360 AAAGACCAAGGACCAGAGCAGGAGCAGACATCAGAAACCTTACAATCTAAGGAAGAAGAT 419
Qy 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
Db 420 GAAGTCAGACAGGCA-----GATAAAGAT 443
Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 444 AATGGAGGTGACTTACAGACTCAAGGCCCATGTGATTCCTAAGTTCGACACAAGTGTG 503
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 504 GATCTACAC-----TATGACAGCCAGAGATGAAATTTATTGTCT 542
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 543 GATGCTTTTAAACCATACCAAGAAACCTTCCAGCCCTCACACCATTTATCTTACATGGAAGA 602
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgGlyLeuValLeuCysTrpAlaGlnGlyGly 215
```



```
QY 932 GlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db 2811 GCAATGAAGTACTGTGTGAAGCTTTAAAGAACTTACTTGTACCTTCAAGAACTGAA 2870
QY 952 LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db 2871 CTAGTGGAGTCTCAACTCACACAGAACTGCTCGAGGACCTGGCTGTATGATCAACA 2930
QY 972 SerArgHisLeuSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAla 991
Db 2931 ACCAAGCACTTAAAGATTTGGATCTTGGTTAAACAAGCCCTGGGTGCAAGAGAGTCA 2990
QY 992 AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db 2991 ACCCTGTGTGAGGACTGAACAAGATGACCTCCCTGAGGAGACTTGGGTGGGGGCA 3050
QY 1012 CysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db 3051 TGTAAGTTGACTTCAATGCTGTGAGGCAATGCTCAATGGCCATCTTTCGAACCTCAC 3110
QY 1032 LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db 3111 CTGAACAGCCTAAACCTGGTGAAGAACTGACTTCAGTACATCGGGGATGTTGAAGCTGTC 3170
QY 1052 SerAlaPheAlaCysProThrSerAsnLeuGlnIleLeuGlyLeuThrProLysTrpGlnTyr 1071
Db 3171 TCTGGTTCATGCTTCAATGCTTCTTAACCTGGGGATAATTTGGCCCTGTGGAAGCAGGATAC 3230
QY 1072 ProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArgValIle 1091
Db 3231 TATGCCGAGTGAGAAGACAGCTGGAGGAAGTTGAGTTGTCAAGCCCGCCTGCTGATT 3290
QY 1092 AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db 3291 GATGGTGAATGCTAGTGTAGTGATGAAGATACCGGAAAC 3329

RESULT 15
AF074018
LOCUS AF074018 3447 bp mRNA linear ROD 10-MAY-2000
DEFINITION Mus musculus MATER protein (Mater) mRNA, complete cds.
ACCESSION AF074018
VERSION AF074018.1 GI:5802697
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Tong, Z. B. and Nelson, L. M.
A mouse gene encoding an oocyte antigen associated with autoimmune
premature ovarian failure
Endocrinology 140 (8), 3720-3726 (1999)
99360614
PUBMED 10433232
REFERENCE 2 (bases 1 to 3447)
Tong, Z.-B. and Nelson, L.M.
Direct Submission
Submitted (24-JUN-1998) Developmental Endocrinology Branch, NICHD,
National Institutes of Health, Building 10, Room 10N262, Bethesda,
MD 20892, USA
FEATURES
Location/Qualifiers
1..3447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
1..3447
/gene="Mater"
9..3344
/gene="Mater"
/note="ooplasm specific protein; OP1"
/codon_start=1
/product="MATER protein"
```

```
/protein_id="AAD51762.1"
/db_xref="GI:5802698"
/translation="MGPPKSKAILKARGLBEEQSKSRMTSPENDSKSIQKQGPPE
QEGTSESTMGPEKSKAILKARGLBEEQSKSRMTSPENDSKSIQKQGPPEQAS
ERMTSPENDSKSIQKQGPPEQSTSETLQSKDEEDTADKONGDLDQDYKAHVAK
FDTSDVLHSDPEMKLLDAFQKTPQHTIILHGRPGVGSALARSIVLGAOGK
LPQKMSFVIFSVREIKWTEKSLAOLIAKECPDSWDLYTKINSQPERLLFVLDGDD
MDSVLQDHTLSRDWDKQPIYILMLSLKALLPQSFLLITTRNTGLEKLSMVS
PLVILVEGLSASRSQVLNENSDRIQVHSLHLENHLPDQCAPSVCSLWVW
OQKLGKRCCTLPQTLGTYATLVFHLTPKPSQSLQSEERQITLVGLCMAAGV
WTMSVFYDDDLKNSYLSKESEIALLPHMILLQVHNSQCVFVSHLSQDPFAALY
VLEGLEWQHCFIENORSIMEVKETDTRLLGMKRFLEGLMKNKILILEVLFPV
VIPTVQKLQHWVSLIAQVNGTSPMDTLDFAFCLFESQDEPFVGALRRFQVWILLI
NQMDLKSYSCLKHKONLKAIRVDRLLSVDNTLELCPVTVQTCRKLWVW
WFCVLSLRNLKELDGLSDISLSQRAKILCLERNSCRIQKLTFSKAVVSWXHL
KLLFNGNKLKYLNGTMMKDDMKLACEALKHPKCSVETRLDSCETIIGYEMIS
TLISITRLKCLSLAKNRVGSVMISLGNALSSMCLLOKLIJLNDGCLTPASCHLVS
ALFSONLTHLCLSNLSLCTEGVQOCLFRLNPECALORLIJLNCNIVDADYGLAMR
LANNTKLTHLSLTMPVPGAMKLLCEALKEPTCYLQLELVDQCLTQNCEDLACMI
TTTHKLSLDLGNALDGGVITLCBGLKQSSSLRLGLGKCLTNSCEALSLAIS
CNPHLSNLNLVKNDFSTGMLKLSAFQCPVSNLGIIGLWKQEYARVRQLEVEFV
KPHVITDGDWYASDEDDRNWKN"
```

ORIGIN

Alignment Scores: 4.54e-192 Length: 3447

Pred. No.: 2688.00 Matches: 554

Score: 66.49% Conservative: 186

Percent Similarity: 49.78% Mismatches: 293

Best Local Similarity: 38.00% Indels: 81

Query Match: 10 Gaps: 11

DB: 11

US-10-066-521-6 (1-1344) x AF074018 (1-3447)

QY 15 GlnTrpCysLeuTyrGluLeuAspLysGluGluPheGlnThrPhelLysGluLeuLys 34

Db 163 CAATGGTCTTCA-GAAAAAGACAGTAAAGCAATCTTGAAGCAGCGTGGATGGAGAG 221

QY 35 LysLysSerSerGluSerThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsn 54

Db 222 GAACAGAGTCAAGAAAGCACA----- 242

QY 55 ValGluCysLeuAlaLeuLeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThr 74

Db 242 ----- 242

QY 75 SerIleSerIlePheGluAsnMetAsnLeuLeuSerGluLysAlaArgAspAsp 94

Db 243 ---ATGTCCTCTTCAAGAAATGTCAGTAGCAGCAATCTTGAAGACAGTGGATCAGAGAA 299

QY 95 MetLysLysIleSerGln----- 100

Db 300 GTGGAACAGCGGTGCAAGAAATAATGACTTCTCCAGAAAAACGACAGTAAATCAATCCAG 359

QY 101 -----AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGln 115

Db 360 AAGACCAAGGACGAGCAGGAGCAGACATCAGAAACCTTACAATCTAAGGAAAGAT 419

QY 116 GluIleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135

Db 420 GAAGTCACAGAGCA-----GATAAGAT 443

QY 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155

Db 444 AATGGAGGTGACTTACAGACTACAGGCGCCATGTGATTGCTAAGTTCGACACAGTGTG 503

QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175

Db 504 GATCTACAC-----TATGACGCCACAGATGAATATTGTCT 542

QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195

Db 543 GATGCTTTTAAACCATTACCAGAAACCTTCCAGCCTCACACCATATCTACATGGAAGA 602

QY 196 SerGlyIleGlySerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
DB 603 CCAGGAGTTGGGAAGTCACTTGGCCAGAAGTATTGTTCTTGGCTGGGCACAGGCTAAA 662
QY 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
DB 663 CTCITTCCTCAAAAATG---TCCTTTGTCATCTCTCTCTGTAGAGAAATAAGGTTGACA 719
QY 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
DB 720 GAGAGAGCAGTTTGGCAGACGTGATTGCTAAGGAGTGTCCAGACTCTCTGGGATCTAGTG 779
QY 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
DB 780 ACAAGATCATGTCTCCCAACCAAGAGACTCTTGTATTGTATAGATGGCTTGGATGATG 839
QY 276 GlySerValLeu---AsnAsnAspThrLysLysCysLysAspTrpAlaGluLysGlnPro 294
DB 840 GACTCTGTCTCCCAACATGATGATGATGACATATCCAGAGACTGGAAGGATGAACAGCCC 899
QY 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
DB 900 ATATACATCTGATGTACAGCTCTCTGAGGAAGGCTCTTTACCTCAGTCTCTTCTCATC 959
QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
DB 960 ATTACCACCAAGAACACAGGCTTAGAAAAACTCAAGTCAATGGTGTGTCTCCCTCTCAT 1019
QY 335 LeuLeuValArgGlyLysSerGlyGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
DB 1020 ATACTGGTTGAAGGAGTGTCTGCATCAAGGAGATCTCAGCTGGTCTCGAGAATCTCTCC 1079
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
DB 1080 AATGAGTCTGATAGATACAGTCTTCCATTCTCTGATAGAAATCACCAGCTGTTTGAC 1139
QY 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
DB 1140 CAATGCCAGGCCCTCTGTGTGTCTCTGCTGCTGCTGTGAGGCTCTACAGCTACAGAGAAA 1199
QY 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
DB 1200 CTGGGAAAGAGATGACCCCTACCTGCCAGACTCTCTACCCGGTTTGTATGCCACGCTGGTG 1259
QY 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgVal 434
DB 1260 TTTCAACAGCTCACCTTGAAAGGCTTCCAGAGCGCTCTCAGTCAGGAAGAACAGATT 1319
QY 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
DB 1320 ACTCTAGTGGGTTTGTGCATGATGCGCAGCTGAAGGAGTGTGGACCATTGAGGTGGTGTTC 1379
QY 455 AspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
DB 1380 TATGATGATGACCTTGAGAACTATAGCCTAAAGGAGTCTGAGATCTTGGCCCTCTTTTAC 1439
QY 475 MetAsnIleLeuLeuProAspSerHis---CysGluGluTyrTyrThrPhePheHisLeu 493
DB 1440 ATGACATCTCTCCAGTTGGCCACAAAGTGAAGGAGTGTATTGTTTCTCCACCTC 1499
QY 494 SerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeu---GluIleGlu 512
DB 1500 AGCCTGCAGGATTTCTTGTCTGCCTTATATTATGTTTGTAGAGGCTGGAGGAATGGAAT 1559
QY 513 ProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAla 532
DB 1560 CAGCATTTTTCG-----TTTCATGAAACCAAGAGGATCATGGAAGTGAAGAACT 1613
QY 533 GlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAsp 552
DB 1614 GAC---GACACTCGCCTCTCGGATGAAGGCTTCTTATTATTGGCCTCATGAACAGGAT 1670

QY 553 ValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLys 572
DB 1671 ATCTTGAAGACTCTGGAGGTTCTGTTTGAATATCCCGTGAATCCCACTGTTGAGCAGAAG 1730
QY 573 LeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThr 592
DB 1731 CTCCAACACTGGGTCTCTCTGATAGCTCAGCAGGCTCAATGGCCAGCCCAATGGRACCC 1790
QY 593 LeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeu 612
DB 1791 CTGGATGCTCTTCTATTGTTCTATTGAGTCTCAGGATGAGAGTTGTTGGCGGGCTCTC 1850
QY 613 AsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuLeuLeuIleAlaSerSer 632
DB 1851 AAACGCTTCCAAGAAGTGTGGCTGTGATTAAACCAGAAGATGGACTTTGAAGGTCTCTTCC 1910
QY 633 PheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePhe 652
DB 1911 TACTGTCTCAAGCAGCTGTGAGAACTTGAAGGCAATCCGGGTGGATATCAGAGACCTCTC 1970
QY 653 ProArgAspGluSerAlaGluAlaCysProValValProLeuTrp---MetArgAspLys 671
DB 1971 TCGGTAGATTAATACTCTCGAGCTGTGCCCTGTTGTTACTCTCCAGGAGACACAATGTAG 2030
QY 672 ThrLeuIleGluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeu 691
DB 2031 CCCCTCTCATGAGTGTGGGAAAACCTTCTGCTGTGCTTGGCAGCCTCCGGAACCTG 2090
QY 692 ArgGlnLeuAspLeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAla 711
DB 2091 AAGAGCTGAGCTTGGGCGCAGCATCTCTGAGTCAACGGGCCCAAGATCTGTGCTCCT 2150
QY 712 LysLeuArgHisProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThr 731
DB 2151 GAGTGGGGAATCAGTCTCGAGATACAGAGCTGACGTTTAAGATGAGCAGGATAGT 2210
QY 732 ProGlyValGlnHisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsn 751
DB 2211 TCTGGCCTGAAACATCTCTGGAAGCTCTTTTAGCAATCAAAACTTTAAAGTACCTCAAT 2270
QY 752 LeuGlyGlyThrHisLeuLysGluLysValArgMetAlaCysGluAlaLeuLysHis 771
DB 2271 CTAGGGAACACTCCCATGAAGGATGATGATGAAGTTAGCTCGAAGCGCTGAAACAT 2330
QY 772 ProLysCysLeuLeuGluSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyr 791
DB 2331 CCAAGTCTCCGTTGGAGACTCTGAGGTGGATTCCTGTGAGTTAACCATCAATGTTAT 2390
QY 792 LeuLysIleSerGlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGly 811
DB 2391 GAGATGATCTCCAGCTTCTTATTTCACACCAGGCTAAAGTGTCTCAGCCTGGCCAAA 2450
QY 812 AsnLysValThrAspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCys 831
DB 2451 AATAGAGTGGGAGTAAAGACATGATATCCCTTGGGAATGCCCTTGATAGTCAATGTGT 2510
QY 832 AlaLeuGlnLysLeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeu 851
DB 2511 CTACTGCAAGTTGATGATCTGGCACTGTGGCTACACCTGCCAGCTGCCACCTCTG 2570
QY 852 AlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSerLeu 871
DB 2571 GTCTCAGCCCTTTTCAGCAACCAAGACTTGACACACCTGTGCTGTCAAAACACAGCCTG 2630
QY 872 GlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGln 891
DB 2631 GGGACTGAAGGAGTGAACAGCTGTGTGCTCTCTGAGGAATCCAGAAATGTGCTCTCCAG 2690
QY 892 ArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAla 911
DB 2691 CGGCTGATCTGATCACTGCAACATTTGATGATGATGCTTATGGCTTCTTGGCAATGAGA 2750
QY 912 LeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsn 931

```
Db      2751  CTGCAACACACAAAGCTGACCCACCTGAGCCTGACCATGAACCCCGTAGGGGATGGT 2810
Qy      932  GlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGlu 951
Db      2811  GCAATGAAGCTACTGTGTGAAGCTTTAAAGGAACCTACTTGTACCTTCAAGAACTGGAA 2870
Qy      952  LeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArg 971
Db      2871  CTAGTGGACTCCCAACTCACACAACTGCTGCGAGGACCTGGCCTGTATGATCACAACA 2930
Qy      972  SerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAla 991
Db      2931  ACCAAGCACTTAAAGATTGGATCTTGGTAAACACGCCCTGGGTGACAAAGGAGTCATA 2990
Qy      992  AlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAla 1011
Db      2991  ACCCTGTGTGAGGGACTGAAGCAAAGTAGCAGCTCCCTGAGGAGACTTGGGTTGGGGCA 3050
Qy      1012  CysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHis 1031
Db      3051  TGTAAAGTTGACTTCCAAATGCTGTGAGGCATGTGCTATGGCCATCTCTGCCAACCCCTC 3110
Qy      1032  LeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCys 1051
Db      3111  CTGAACAGCCTAAACCTGGTGAAGAAATGACTTCAATACATCGGGGATGTTGAAGCTGTC 3170
Qy      1052  SerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTyr 1071
Db      3171  TCTGCGTTCCAATGCCCTGTCTCTAACTGGGGATAATTGGCCTGTGGAGCAGGAGTAC 3230
Qy      1072  ProValGlnIleArgLysLeuLeuGluGluValGlnLeuLysProArgValValIle 1091
Db      3231  TATGCCCGAGTGAGAACAGACAGCTGGAGGAAGTTGAGTTGTCAAGCCCCACGTCGTGATT 3290
Qy      1092  AspGlySerTrpHisSerPheAspGluAspArgHis 1104
Db      3291  GATGGTGATTGGTATGTAGTGTAGTAGAGATGACCCGAAAC 3329
```

Search completed: July 16, 2004, 04:26:17
Job time : 12554 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 09:07:16 ; Search time 7396 Seconds
(without alignments)
5426.551 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 1344
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSNLGAAGLEGLVS 1344

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US10066521/runat_13072004_122317_11548/app_query.fasta_1.1543
-DB=EST -Qfmt=fastap -SUFFIX=Oligo.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn_1.1.5759 @runat_13072004_122317_11548 -NCPU=6 -ICPU=3
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	226	16.8	791	14	CK000513	CK000513 AGENCOURT
2	107	8.0	693	13	BU634350	BU634350 UI-H-FL1-
3	99	7.4	684	13	BU630481	BU630481 UI-H-FL0-
4	85	6.3	642	13	BU618831	BU618831 UI-H-FL1-
5	38	2.8	526	28	AQ542616	AQ542616 RPCI-11-3
6	20	1.5	523	28	AQ428788	AQ428788 CJTBI-E1-
7	20	1.5	814	14	CB228956	CB228956 AGENCOURT
8	16	1.2	743	29	AG117567	AG117567 Pan trogl
9	12	0.9	706	10	BB624558	BB624558 BB624558
10	12	0.9	1913	11	AK016782	AK016782 Mus muscu
11	12	0.9	3475	11	AK087774	AK087774 Mus muscu
12	11	0.8	394	14	CF916020	CF916020 B0988G02-
13	11	0.8	511	14	CF173704	CF173704 B0927G04-
14	11	0.8	533	14	CA559680	CA559680 K0261H12-
15	11	0.8	546	14	CA561388	CA561388 K0286C11-
16	11	0.8	559	14	CA561556	CA561556 K0288H05-
17	11	0.8	583	14	CA559979	CA559979 K0266B01-
18	11	0.8	589	14	CA559886	CA559886 K0264H01-
19	11	0.8	593	14	CA558845	CA558845 K0250F09-
20	11	0.8	599	14	CF915731	CF915731 B0984B03-
21	11	0.8	636	14	CF914959	CF914959 B0971H02-
22	11	0.8	643	14	CF913803	CF913803 B0954D05-
23	11	0.8	664	29	AG067278	AG067278 P8n tX0g1
24	11	0.8	1119	12	BM454123	BM454123 AGENCOURT
25	10	0.7	250	28	BZ170794	BZ170794 CH230-247
26	10	0.7	355	10	BF772019	BF772019 IL5-IT002
27	10	0.7	466	12	BM192717	BM192717 da120H03
28	10	0.7	515	10	BE726276	BE726276 89409D08
29	10	0.7	550	28	BZ302590	BZ302590 K01924.P1
30	10	0.7	754	14	CA817466	CA817466 CAL2E1203
31	9	0.7	172	9	AV421688	AV421688 AV421688
32	9	0.7	192	11	CNS090HH	BX036705 Single re
33	9	0.7	199	12	BJ242372	BJ242372 BJ242372
34	9	0.7	206	10	AW686846	AW686846 NF003B01R
35	9	0.7	236	12	BJ240882	BJ240882 BJ240882
36	9	0.7	238	28	AZ343507	AZ343507 IM0077105
37	9	0.7	263	12	BJ235665	BJ235665 BJ235665
38	9	0.7	264	9	AA231647	AA231647 CD0328.R
39	9	0.7	300	12	BJ262094	BJ262094 BJ262094
40	9	0.7	300	28	B49653	B49653 CIT-HSP-242
41	9	0.7	301	12	BJ308866	BJ308866 BJ308866
42	9	0.7	302	9	AA853292	AA853292 NHTBca604
43	9	0.7	311	13	BU992299	BU992299 HD09M01r
44	9	0.7	319	12	BJ228111	BJ228111 BJ228111
45	9	0.7	323	10	BB123682	BB123682 BB123682

ALIGNMENTS

RESULT 1
CK000513 791 bp mRNA linear EST 26-NOV-2003
LOCUS AGENCOURT_16368905 NIH_MGC_221 Homo sapiens cDNA clone
DEFINITION IMAGE:30708637 5', mRNA sequence.
CK000513
ACCESSION CK000513.1 GI:38526547
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 791)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: ND41075 row: c column: 14
 High quality sequence stop: 688.
 Location/Qualifiers
 1..791

1..791

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30708637"

/lab_host="DH10B Tona"

/clone_lib="NIH MGC 221"

/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;

Site 2: NotI; Library is oligo-dT primed and directionally

cloned. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. Average insert size

4-5Kb. Adaptors 5' (AATCGGACGAGG)3' and 5'd

(CCTCGTCCG)3'. 3' linker sequence - GCGGCCGCTGAGAGCC T18.

Sequencing primers 3' end: T3 promoter primer 5'd

(ATTAACTCCTCAATAAGGGA)3'. 5' End: T7 promoter primer 5'd

(TAATAGCCTCACTATAGG)3'. Library was constructed in the

laboratory of M. Bento Soares. Note: this is a NIH_MGC

Library"

ORIGIN

Alignment Scores:

Pred. No.: 3 24e-227 Length: 791
 Score: 226.00 Matches: 242
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 16.82% Indels: 1
 DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CK000513 (1-791)

Qy 190 ValValLeuHisGlyLysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeu 209
 Db 8 GTGGTTCTGCAGCGAAAGTCAGGAATGGGAATCGCTCTAGCCAGAGGATGTCGTG 67
 Qy 210 CysTrpAlaGlnGlyLeuTyGlnGlyMetPheSerTyValPheLeuProVal 229
 Db 68 TGCTGGGCGCAAGGTGGACTCTACAGGGAATGTTCTCTACGTCTTCTCTCCCGGTT 127
 Qy 230 ArgGluMetGlnArgLysLysSerValThrGluPheLeuSerArgGluTrpPro 249
 Db 128 AGAGAGATGCAGCGAAG 187
 Qy 250 AspSerGlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleIle 269
 Db 188 GACTCCCGAGGTCCTCGGTGACGAGATCATGTCTCCGACCAAGAGCTGTGTTTCATCAT 247
 Qy 270 AspGlyPheAspLeuGlySerValLeuAsnAsnAspThrLysLeuLysLysAspTrp 289
 Db 248 GACGGTTTCGATGACCTGGGCTCTGTCTCTCAACATGACACAAAGCTCTGCAAGACTGG 307
 Qy 290 AlaGluLysGlnProProPheThrLeuIleArgSerLeuLeuArgLysValLeuPro 309

Db 308 GCTGAGAGAGAGCTCGTTCACCTCATACGAGCTCTGCTGAGGAAGTCTCTGCTCCCT 367
 Qy 310 GluSerPheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluVal 329
 Db 368 GAGTCTCTCTGATCGTACCGTCAGAGAGCTGGGACAGAGAAGCTCAAGTCAGAGGTC 427
 Qy 330 ValSerProArgTyTrileuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeu 349
 Db 428 GTGTCTCTCCCGTACCTGTAGTGTAGAGGAATCTCCGGGGAAACAAAGAAATCCACTTGTCTC 487
 Qy 350 LeuGluArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsn 369
 Db 488 CTTGAGCGCGGATTTGGTGGCATCAGAGACACAGAGGTGGTGGCATCATGAACAAC 547
 Qy 370 ArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLeuLeuCysValAlaLeu 389
 Db 548 CGTGAGCTGCTCGACCATGTCAGGTGCCCGCGTGGGCTCTCTCATCTGCTGGGCGCTG 607
 Qy 390 GlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeu 409
 Db 608 CAGCTGCAGGACGTGGTGGGGAGAGCGTGCCTCCCTTCAACCAACGCTTCACAGGCTG 667
 Qy 410 HisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsn 429
 Db 668 CAGCGCGCTTTGTGTTTAT-CAGCTACCCCTCAGAGCGTGGTGGGCGCTGTCTCAAT 726
 Qy 430 LeuGluGlu 432
 Db 727 CTGAGAGAA 735

RESULT 2

BU634350/c

LOCUS

DEFINITION

UI-H-FLI-bgx-m-16-0-UI.s1 NCI CGAP FLI Homo sapiens cDNA clone

UI-H-FLI-bgx-m-16-0-UI 3', mRNA sequence.

ACCESSION

BU634350

VERSION

BU634350.1 GI:23301605

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 693)

NCI-Cgap <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: James Martin

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1..693

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-FLI-bgx-m-16-0-UI"

/tissue_type="Cell lines"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP FLI"

/notes="Organ: Chondrosarcoma; Vector: pT7T3-Pac

(Pharmacia) with a modified polylinker; Site 1: EcoR I;

Site 2: Not I; NCI CGAP FLI is a normalized cDNA library

derived from a pool of mRNA obtained from 4 cell lines

from grade III chondrosarcoma tissues. The library was

constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGCTG. The cell lines were provided by Dr. James Martin from the University of Iowa.

TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLI
TAG_SEQ=GAGTCCGCTG

ORIGIN

Alignment Scores:
Pred. No.: 1,586-101 Length: 693
Score: 107.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.96% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x BU634350 (1-693)

Qy 997 LeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSer 1016
Db 642 CTGAAGCAAAAGACAGTGTCTGACGACTCGGGTTGAAGGCATGTGGACTCTCT 583
Qy 1017 AspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsn 1036
Db 582 GATTGCTGTGAGGCACTCTCTTGGCCCTTCTGCAACGGCATCTGACCATCTAAAC 523
Qy 1037 LeuValGlnAsnAsnPheserProLysGlyMetMetLysLeuLysSerAlaPheAlaCys 1056
Db 522 CTGGTGCAGATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCCTTTGCTGT 463
Qy 1057 ProThrSerAsnLeuGlnIleileGlyLeuThrPlySerTrpGlnTyrProValGlnIleArg 1076
Db 462 CCACGCTTAACTTACAGATAATTGGGCTGTGGAATGCGAGTACCCCTGTGCAATAAGG 403
Qy 1077 LysLeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHis 1096
Db 402 AGCTGCTCGAGGAAGTGCAGTACTCAAGCCCCGAGTCGTGAATGACGGTAGTTGGCAT 343
Qy 1097 SerPheAspGluAspAspArg 1103
Db 342 TCTTTTGTGATGAAGATGACCGG 322

RESULT 3

BU630481/c
LOCUS
DEFINITION
UI-H-FLI-bdk-c-22-0-UI-s1 NCI CGAP FLI0 Homo sapiens cDNA clone
UI-H-FLI-bdk-c-22-0-UI 3', mRNA sequence.

ACCESSION
BU630481

VERSION
BU630481.1

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 684)

AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

COMMENT
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: James Martin

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FLI-bdk-c-22-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="NCI CGAP FLI0"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FLI0 is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGCTG. The cell line was provided by Dr. James Martin from University of Iowa."
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLI
TAG_SEQ=GAGTCCGCTG

ORIGIN

Alignment Scores:
Pred. No.: 4,41e-93 Length: 684
Score: 99.00 Matches: 99
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.37% Indels: 0
DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x BU630481 (1-684)

Qy 990 ValAlaAlaLeuLysGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeu 1009
Db 656 GTTCTGCGCTGTGCGAGGAGCTGACCAAAAGACAGTGTCTGACGAGACTCGGGTTG 597
Qy 1010 LysAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsn 1029
Db 596 AAGGCATGTGAGCTGACTTCTGATTGCTGTGAGGCACTCTCTTGGCCCTTTCCTGCAAC 537
Qy 1030 ArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheserProLysGlyMetMetLys 1049
Db 536 CGGCATCTGACCACTTAAACCTGGTGCAGATAACTTCAGTCCCAAGGAATGATGAAG 477
Qy 1050 LeuCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleileGlyLeuThrLysTrp 1069
Db 476 CTGTGTCGGCCTTGTGCTGTCCACGCTTACACTACAGATAATGCGCTGTGGAATGG 417
Qy 1070 GlnTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLysProArg 1088
Db 416 CAGTACCTGTGCAAAATAAGGAAGCTGCTGAGGAGAGTGCAGCTACTCAAGCCCCGA 360

RESULT 4

BU618831/c

LOCUS

DEFINITION

UI-H-FH1-bfl-i-19-0-UI-s1 NCI CGAP FH1 Homo sapiens cDNA clone

UI-H-FH1-bfl-i-19-0-UI 3', mRNA sequence.

ACCESSION

BU618831

VERSION

BU618831.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 642)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 source
 1..642
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FH1-bf1-i-19-0-UI"
 /tissue type="Cell Line"
 /dev stages="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_FH1"
 /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP_FH1 is a normalized CDNA library
 obtained from a cell line derived from grade I
 chondrosarcoma tissue. The library was constructed and
 normalized according to Bonaldo, Lennon and Soares, Genome
 Research, 6:791-806, 1996. First strand cDNA synthesis was
 primed with an oligo-dT primer containing a Not I site.
 Double stranded cDNA was ligated to an EcoR I adaptor,
 digested with Not I, and cloned directionally into
 pTT73-Pac vector. The oligonucleotide used to prime the
 synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AGAATCCGGC. The cell line was provided by Dr. James Martin
 from the University of Iowa.
 TAG_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
 Chondrosarcoma
 TAG_LIB=UI-H-FH1
 TAG_SEQ=AGAATCCGGC"

ORIGIN
 Alignment Scores:
 Pred. No.: 2.56e-78 Length: 642
 Score: 85.00 Matches: 85
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.32% Indels: 0
 DB: 13 Gaps: 0

US-10-066-521-6 (1-1344) x B0618831 (1-642)

Qy 1004 LeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysGluAlaLeuSer 1023
 |||
 Db 614 CTGACGAGACTCGGGTGAAGGCATGTGGACTGCTTCTGATGCTGTGAGGCACCTCTCC 555
 |||

Qy 1024 LeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSer 1043
 |||
 Db 554 TTGGCCCTTTCTCTCAACCGGCATCTGCACAGTCTAAACCTGGTGCGAGATAACTTCAGT 495
 |||

Qy 1044 ProLysGlyMetMetLysLysCysSerAlaPheAlaCysProThrSerAsnLeuGlnIle 1063
 |||
 Db 494 CCCAAGGAATGATGAAGCTGTGTGGCCCTTTCCTGTGCCAGCTTAACCTTACAGATA 435
 |||

Qy 1064 IleGlyLeuTrpLysTrpGlnTy-ProValGlnIleArgLysLeuLeuGluValGln 1083
 |||

Db 434 ATTGGGCTGTGGAAATGGCAGTACCTCTGTCAAATAGGAAGTCTGCTGGAGGAGTGCAG 375
 |||

Qy 1084 LeuLeuLysProArg 1088
 |||

Db 374 CTACTCAAGCCCGA 360
 |||

RESULT 5
 AQ542616/c
 LOCUS
 DEFINITION
 RPCI-11-345A9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-345A9,
 genomic survey sequence.
 ACCESSION
 AQ542616
 VERSION
 AQ542616.1 GI:4873072
 KEYWORDS
 GSS.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 526)
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
 Venter,J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished (1997)
 Other_GSSs: RPCI-11-345A9.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
 Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..526
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7632104"
 /db_xref="taxon:9606"
 /clone="RPCI-11-345A9"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN
 Alignment Scores:
 Pred. No.: 9.52e-29 Length: 526
 Score: 38.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.83% Indels: 0
 DB: 28 Gaps: 0

US-10-066-521-6 (1-1344) x AQ542616 (1-526)

Qy 840 AspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArg 859
 |||

Db 300 GACTGTGGCATACAGCCACGGTTCAGAGTCTGGCCCTCAGCCCTCTGACCAACCCG 241
 |||

Qy 860 SerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
 |||

Db 240 AGCTTGACACACTGTGCTCTATCCCAACACAGCCTGGGGAACGAAGGTGTAAT 187
 |||


```

RESULT 6
AQ428788      523 bp  DNA  linear  GSS 24-MAR-1999
LOCUS
DEFINITION  CITBI-E1-2560117.TF CITBI-E1 Homo sapiens genomic clone 2560117,
              genomic survey sequence.
ACCESSION   AQ428788
VERSION     AQ428788.1  GI:4496554
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 523)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
              Venter,J.C.
TITLE       Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
              Map Building
JOURNAL     Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2560117.TR
              Contact: Shaying Zhao, William Nierman, Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: hbe@tigr.org
              Clones are available from Research Genetics (info@resgen.com). BAC
              end search page:
              http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
              Seq primer: M13-21
              Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..523
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2560117"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"
ORIGIN
Alignment Scores:
Pred. No.:          9, 93e-10      Length:          523
Score:              20.00          Matches:          20
Percent Similarity: 100.00%        Conservative:    0
Best Local Similarity: 100.00%      Mismatches:     0
Query Match:        1.49%          Indels:         0
DB:                 28             Gaps:            0

US-10-066-521-6 (1-1344) x AQ428788 (1-523)
Qy  116  GluIleSerGlnAlaMeGluGlnGlnGluGlyAlaThrAlaAgluThrGluGluGlnGly 135
Db  22  GAATTCACAGCTATGCAACAAGAGTGCCACAGCAGCAGACAGACAGAACAGGT 81

RESULT 7
CB228956      814 bp  mRNA  linear  EST 10-FEB-2003
LOCUS
DEFINITION  AGENCOURT 11501187 NICHD Rh Ovl Macaca mulatta cDNA clone
              IMAGE:6884760 5', mRNA sequence.
ACCESSION   CB228956
VERSION     CB228956.1  GI:28280534
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM    Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE   1 (bases 1 to 814)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

```

```

TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. Eliot Spindel
              cDNA Library Preparation: CLONTECH
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA sequencing by: Agencourt Bioscience Corporation
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LUCM3135 row: b column: 23
              High quality sequence stop: 235.
              Location/Qualifiers
FEATURES             source
                     1..814
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:6884760"
                     /tissue_type="Ovary"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD Rh Ovl"
                     /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
                     Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.0-4.0 kb. Tissue pooled from
                     pre-pubertal, post pubertal sn menopausal monkeys.
                     Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN
Alignment Scores:
Pred. No.:          1, 62e-09      Length:          814
Score:              20.00          Matches:          20
Percent Similarity: 100.00%        Conservative:    0
Best Local Similarity: 100.00%      Mismatches:     0
Query Match:        1.49%          Indels:         0
DB:                 14             Gaps:            0

US-10-066-521-6 (1-1344) x CB228956 (1-814)
Qy  743  AlaaSnArghenLeuArgSerLeuAsnLeuGlyThrHisLeuLyvGluGluAspVal 762
Db  78  GCCAACCGTAACCTAAGATCCCTCAATTGGAGGACCCACCCTGAAGGAGGAGGTGA 137

RESULT 8
AG117567/c     743 bp  DNA  linear  GSS 03-NOV-2001
LOCUS
DEFINITION  Pan troglodytes DNA, clone: PTB-125H04.F, genomic survey sequence.
ACCESSION   AG117567
VERSION     AG117567.1  GI:16738086
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE     1
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
              Totoki,Y., Watanabe,H. and Sakaki,Y.
              BAC end sequences of Library PTB
              Unpublished
              2 (bases 1 to 743)
              Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
              Totoki,Y., Watanabe,H. and Sakaki,Y.
              Direct Submission
              Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
              1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              (E-mail:chihirobes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
              Tel:81-45-503-9111, Fax:81-45-503-9170)
              Clones are derived from the chimpanzee BAC library PTB This BAC end
              was generated during the R&D process and may have higher chance of
              clone tracking errors.
              PRIMERS

```



```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
  Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
  Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
  Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
  Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
  Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4 The RIKEN Genome Exploration Research Group Phase II Team and the
  FANTOM Consortium.
  Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
  Group Phase I & II Team.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1913)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
  Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
  Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
  Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
  Kaekawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
  Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
  Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
  Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
  Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
  Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
  Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
  Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
  Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
  Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
  Fax: 81-45-503-9216]
Please visit our web site (http://genome.gsc.riken.go.jp/) for
  further details.
cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues. First strand cDNA was primed with a primer
  [5'-GAGAGAGAGAGATCCAGAGCTCTTTTITTTTITTTN 3'], cDNA was
  prepared by using trehalose thermo-activated reverse transcriptase
  and subsequently enriched for full-length by cap-trapper. Second
  strand cDNA was prepared with the primer adapter of sequence [5'
  GAGAGAGAGATTCGAGTAAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved
  with BamHI and XhoI. cDNA of size comprised longer than 7 kb was
  selected before cloning. Vector: a modified pBluescript KS(+) after
  bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
  end: BamHI. Host: DH10B.
  Location/Qualifiers
    source
      1..1913
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="CS7BL/6J"
      /db_xref="FANTOM DB:4933411P06"
      /db_xref="MGI:1903594"
      /db_xref="taxon:10090"
      /clone="4933411P06"
      /sex="male"
      /tissue_type="testis"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="adult"
      735..1544
      /note="unnamed protein product; maternal effect gene
      (MGD|MGI:1345193)"
      putative"
      /codon_start=1
      /protein_id="BAB30427.1"
      /db_xref="GI:12855703"
      /translation="MKDDMKLACEALEHPKCALELRDLSDCGITPASCHLYSALES
      NQNLHLCLNNSLGHTEGVQQLCQFLNPECAQLRLINHCNIVDDAYGFLARLANN
      TKUTHLSLTWNVPVGDGAMKLLCALKEPTCYLKLKLVDCQNMQNCEDLACM1TTTK
      HLKSLDIGNNALGDKGVITLCEGLKQSSSLRLGLGACELTSCCCELSLAI5CNPH
      LNSLNLMDQDFTSGKLLCSAFCPCVSNLGIISGNLGSFALTFTALRRN"
      1896..1891
      /note="putative"
      1913
      /note="putative"
    polyA_signal
    polyA_site
    ORIGIN
  Alignment Scores:
    Pred. No.: 1..19 Length: 1913
    Score: 12.00 Matches: 12
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 0.89% Indels: 0
    DB: 11 Gaps: 0
  US-10-066-521-6 (1-1344) x AK016782 (1-1913)
  Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGly 872
  Db 876 TTGACACACCTGTGCTGTCAACACACAGCCTGGG 911
  RESULT 11
  AK087774
  LOCUS
  DEFINITION
  Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
  full-length enriched library, clone:E330019F16 product:mouse WATER
  protein (maternal-antigen-that-embryos-require) protein, full
  insert sequence.
  AK087774
  VERSION
  AK087774.1 GI:26104500
  KEYWORDS
  HTC; CAP trapper.
  SOURCE
  Mus musculus (house mouse)
  ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
  REFERENCE
  1 Carninci, P. and Hayashizaki, Y.
  TITLE
  High-efficiency full-length cDNA cloning
  JOURNAL
  Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE
  PUBMED
  10349636
  2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes
  Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
  MEDLINE
  PUBMED
  11042159
  3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
  Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
  Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
  Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
  Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
  Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multipillar sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
  MEDLINE
  PUBMED
  11076861
  4 The RIKEN Genome Exploration Research Group Phase II Team and the
  FANTOM Consortium.
  Functional annotation of a full-length mouse cDNA collection
  Nature 409, 685-690 (2001)
  JOURNAL
  MEDLINE
  PUBMED
  11076861
  5 The FANTOM Consortium and the RIKEN Genome Exploration Research
  Group Phase I & II Team.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  Nature 420, 563-573 (2002)
  JOURNAL
  MEDLINE
  PUBMED
  11076861
  6 (bases 1 to 1913)
  Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
  Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
  Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
  Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
  Kaekawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
  Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
  Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
  Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
  Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
  Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
  Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
  Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
  Direct Submission
  Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
  Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
  Fax: 81-45-503-9216]
  Please visit our web site (http://genome.gsc.riken.go.jp/) for
  further details.
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues. First strand cDNA was primed with a primer
  [5'-GAGAGAGAGATCCAGAGCTCTTTTITTTTITTTN 3'], cDNA was
  prepared by using trehalose thermo-activated reverse transcriptase
  and subsequently enriched for full-length by cap-trapper. Second
  strand cDNA was prepared with the primer adapter of sequence [5'
  GAGAGAGAGATTCGAGTAAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved
  with BamHI and XhoI. cDNA of size comprised longer than 7 kb was
  selected before cloning. Vector: a modified pBluescript KS(+) after

```

```

MEDLINE      20499374
PUBMED       11042159
REFERENCE    3
AUTHORS      Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
              Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
              Sumi,N., Ishii,Y., Nakamura,S., Kizama,M., Nishine,T., Harada,A.,
              Yamamoto,R., Matsuoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
              Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,
              Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
              Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
              RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
              Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL      20530913
MEDLINE      11076861
PUBMED
REFERENCE    4
AUTHORS      The RIKEN Genome Exploration Research Group Phase II Team and the
              FANTOM Consortium.
              Functional annotation of a full-length mouse cDNA collection
              Nature 409, 585-690 (2001)
JOURNAL      5
PUBMED
REFERENCE
AUTHORS      The FANTOM Consortium and the RIKEN Genome Exploration Research
              Group Phase I & II Team.
              Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
              Nature 420, 563-573 (2002)
JOURNAL      6
PUBMED
REFERENCE
AUTHORS      Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
              Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
              Hayashida,K., Hayatsu,N., Hiramoto,K., Hitaoka,T., Hirozane,T.,
              Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
              Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
              Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
              Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
              Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
              Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
              Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
              Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
              Muramatsu,M. and Hayashizaki,Y.
              Direct Submission
              Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
              RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
              Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp,
              URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
              Fax:81-45-503-9216]
              cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues.
              Please visit our web site for further details.
              URL:http://genome.gsc.riken.go.jp/
              URL:http://fantom.gsc.riken.go.jp/.
FEATURES
  Location/Qualifiers
    1..3475
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="FANTOM,DB:E330019F16"
      /db_xref="WGI:2426482"
      /db_xref="WGI:2426482"
      /db_xref="taxon:10090"
      /clone="E330019F16"
      /sex="female"
      /tissue_type="ovary"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="2 days pregnant adult"
    77..3366
      /note="mouse MATER protein
      (maternal-antigen-that-embryos-require) protein
      putative"
      3452..3457
polyA_signal

```

```

polyA_site    /note="putative"
              3475
              /note="putative"
ORIGIN
Alignment Scores:
Pred. No.:    2.3      Length:      3475
Score:        12.00    Matches:    12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB:           11      Gaps: 0
US-10-066-521-6 (1-1344) x AK087774 (1-3475)
QY      861 LeuthHisLeuCysLeuSerAsnSerLeuGly 872
Db      2620 TTGACACACCTGTGCTGTCAAAACACAGCCTGGG 2655
RESULT 12
LOCUS     CF916020
DEFINITION Mus musculus (house mouse)
ACCESSION CF916020
VERSION    CF916020.1 GI:38187222
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 394)
AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
JOURNAL    Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE    21429098
PUBMED     11544199
COMMENT    Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: B0988 row: G column: 02
            Seq primer: M13 Reverse
            High quality sequence stop: 394
            POLYA=No.
FEATURES
  Location/Qualifiers
    1..394
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="niaEST:B0988G02-5"
      /db_xref="taxon:10090"
      /clone="NIA:B0988G02 IMAGE:30480745"
      /dev_stage="Unfertilized Egg"
      /lab_host="DH10B"
      /clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
      1)"
      /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
      Site 2: NotI; Mouse cDNA project by the Laboratory of
      Genetics, National Institute on Aging (NIA), Intramural
      Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
      This is a long-transcript enriched cDNA library [Ref.
      Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]]. Total
      RNAs were extracted from a pool of 1488 unfertilized eggs.
      Double-stranded cDNAs were synthesized with an Oligo(dT)
      primer [Invitrogen:
      5'-pGACTAGTTCATGATCGAGCGCGCCGCGCCCTTTTCTTTT-3'],
      treated with T4 DNA polymerase, and purified by
      ethanol-precipitation. The cDNAs were ligated to
      Lone-linker LL-Sal4, purified by phenol/chloroform, and

```

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 2.35 Length: 394
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CF916020 (1-394)

Qy 500 AlaAlaLeuTyTyValLeuGluGlyLeuGlu 510

Db 140 GCTGCTTATATTATGTTTATAGAGGGCTGGAG 172

RESULT 13

CF173704

LOCUS CF173704 511 bp mRNA linear EST 25-JUL-2003
DEFINITION B0927G04-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
musculus cDNA clone NIA:B0927G04 IMAGE:30474891 5', mRNA sequence.

ACCESSION CF173704

VERSION CF173704.1 GI:33283253

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE 21429098

PUBMED 11544199

COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: B0927 row: G column: 04

Seq primer: M13 Reverse

High quality sequence stop: 511

POLYA=No.

FEATURES

source

1. 511 Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0927G04-5"
/db_xref="taxon:10090"
/clone="NIA:B0927G04 IMAGE:30474891"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: Sali;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cdna).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total

RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):

5'-pGACTAGTCTAGATCGAGCGGCCCTTTT-TTTT-3'],

treated with T4 DNA polymerase, and purified by

ethanol-precipitation. The cDNAs were ligated to

Lone-linker LL-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with Sali and NotI enzymes

and cloned into Sali/NotI site of pCMV-SPORT6 plasmid

vector. The DH10B E. coli host was transformed with the

ligation mixture by the standard chemical method. The

average insert size is about 2.5 kb. The library was

constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 3.14 Length: 511
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CF173704 (1-511)

Qy 500 AlaAlaLeuTyTyValLeuGluGlyLeuGlu 510

Db 271 GCTGCTTATATTATGTTTATAGAGGGCTGGAG 303

RESULT 14

CA559680

LOCUS

DEFINITION K0261H12-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
musculus cDNA clone NIA:K0261H12 IMAGE:30050591 5', mRNA sequence.

ACCESSION CA559680

VERSION CA559680.1 GI:25104279

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Staggs, C.A.,
Martin, P., Alba, K., Tanaka, T. and Ko, M.S.H.

SYSTEMATIC ANALYSES OF NIA Mouse Unfertilized Egg cDNA Library

(Long)

JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0261 row: H column: 12
Seq primer: M13 Reverse
High quality sequence stop: 533
POLYA=No.

FEATURES

source

1. 533 Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:K0261H12-5N"
/db_xref="taxon:10090"
/clone="NIA:K0261H12 IMAGE:30050591"
/tissue_type="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: Sali; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
 5'-pGATGATCTAGATCGGAGCGCGCCCTTTT-3',
 treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Alignment Scores:
 Pred. No.: 3.29 Length: 533
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CA559680 (1-533)

Qy 500 AlaAlaLeuTyrValLeuGluGlyLeuGlu 510

Db 427 GCTGCTTATATTATGTTTTAGAGGCTGGAG 459

RESULT 15

CA561388 546 bp mRNA linear EST 19-NOV-2002
 LOCUS K0286C11-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
 DEFINITION musculus cDNA clone NIA:K0286C11 IMAGE:30052930 5', mRNA sequence.

CA561388
 VERSION CA561388.1 GI:25106043
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 546)
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Staggs, C.A.,
 Martin, P., Aliba, K., Tanaka, T. and Ko, M.S.H.
 Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library
 (Long)

Unpublished (2001)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: K0286 row: C column: 11

Seq primer: M13 Reverse

High quality sequence stop: 546

POLYA=No.

FEATURES

source

Location/Qualifiers
 1..546
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="C57BL/6J"
 /db_xref="niaEST:K0286C11-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0286C11 IMAGE:30052930"

/tissue_type="Unfertilized Egg"

/lab_host="DH10B"

/clone_lib="NIA Mouse Unfertilized Egg cDNA Library

(Long)"

/Notes="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were
 extracted from a pool of 1488 unfertilized eggs.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen]:

5'-pGATGATCTAGATCGGAGCGCGCCCTTTT-3',

treated with T4 DNA polymerase, and purified by

ethanol-precipitation. The cDNAs were ligated to

Lone-linker IL-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation

mixture by the standard chemical method. The average

insert size is about 2.5 kb. The library was constructed

by Yulan Piao (NIA)."

ORIGIN

Alignment Scores:
 Pred. No.: 3.38 Length: 546
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 14 Gaps: 0

US-10-066-521-6 (1-1344) x CA561388 (1-546)

Qy 500 AlaAlaLeuTyrValLeuGluGlyLeuGlu 510

Db 427 GCTGCTTATATTATGTTTTAGAGGCTGGAG 459

Search completed: July 16, 2004, 15:28:15

Job time : 7413 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 06:26:49 ; Search time 12085 Seconds
(without alignments)
4820.276 Million cell updates/sec

Title: US-10-066-521-6
Perfect score: 1344
Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGLVS 1344

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10066521/runat_13072004_122317_11534/app_query.fasta_1.1543
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn 1 1 8359 @runat_13072004_122317_11534 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sv.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.ma.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

Result No.	Score	Query Match %	Length	DB ID	Description
1	1006	74.9	3489	6	AX478549 Sequence
2	1006	74.9	3885	9	AY154460 Homo sapi
3	1006	74.9	3926	6	AX704821 Sequence
4	988	73.5	3830	6	AX704823 Sequence
5	968	72.0	5859	6	AX459881 Sequence
6	968	72.0	6939	6	AX459873 Sequence
7	968	72.0	6939	6	AX459891 Sequence
8	616	45.8	3885	9	AY054986 Homo sapi
9	533	39.7	157141	9	AC011470 Homo sapi
10	533	39.7	193609	9	AC024580 Homo sapi
11	385	28.6	1157	6	AX427588 Sequence
12	300	22.3	167509	2	AC012107 Homo sapi
13	300	22.3	167509	2	AX427590 Sequence
14	196	14.6	1075	6	AC012107 Homo sapi
15	169	12.6	167509	2	AC012107 Homo sapi
16	75	5.6	164824	2	AC023887 Homo sapi
17	57	4.2	164824	2	AC023887 Homo sapi
18	55	4.1	1705	9	BC040925 Homo sapi
19	12	0.9	171	10	FI43559S11 Mus muscu
20	12	0.9	3108	6	AX417214 Sequence
21	12	0.9	3221	9	AY116207 Homo sapi
22	12	0.9	3300	6	AX459869 Sequence
23	12	0.9	3303	10	AY329488 Mus muscu
24	12	0.9	3303	10	AY329491 Mus muscu
25	12	0.9	3372	10	AY329485 Mus muscu
26	12	0.9	3375	10	AY329489 Mus muscu
27	12	0.9	3395	9	AY116206 Homo sapi
28	12	0.9	3405	10	AY329487 Mus muscu
29	12	0.9	3432	10	AY329484 Mus muscu
30	12	0.9	3447	6	AX427592 Sequence
31	12	0.9	3447	10	AF074018 Mus muscu
32	12	0.9	3453	10	AY329486 Mus muscu
33	12	0.9	3454	10	AY329490 Mus muscu
34	12	0.9	3466	6	AX833594 Sequence
35	12	0.9	3466	9	AK095460 Homo sapi
36	12	0.9	3470	10	AY196362 Mus muscu
37	12	0.9	3480	10	AY196361 Mus muscu
38	12	0.9	3507	9	BC028069 Homo sapi
39	12	0.9	3534	10	BC053384 Mus muscu
40	12	0.9	3563	9	AY116205 Homo sapi
41	12	0.9	3731	9	AY116204 Homo sapi
42	12	0.9	3827	9	AY095146 Homo sapi
43	12	0.9	3827	9	AY154467 Homo sapi
44	12	0.9	4931	6	AX684291 Sequence
45	12	0.9	67802	2	AC101178 Mus muscu

ALIGNMENTS

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

AX478549
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO248362.
ACCESSION AX478549
VERSION AX478549.1 GI:22217318
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Ramkumar,J. and Arvizu,C.
TITLE Embryogenesis associated proteins
JOURNAL Patent: WO 0248362-A 3 20-JUN-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1. 3489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474830CB1"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3489
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX478549 (1-3489)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaGluThrGluGluGlnGlu 117
Db 454 ATTTCAAGCTATGGAACAGAGGTGCCACAGCAGCAGACAGACAGAAAGAAATT 513
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaGluThrGluGlnGlyHisGly 137
Db 514 TCACAAAGCTATGGAACAGAGGTGCCACAGCAGCAGACAGAAAGACATGGA 573
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 574 GGTGACACATGGACTACAGAGTCACGTGATGACCAATTTCGCTGAGGAGGAGTGA 633
Qy 158 ArgArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTGAACACACTGCTGCTGACTGCGCGGGAATGCAACGTTGGCTGGTCT 693
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 694 TTTGATTTCAGACCGGTGGGGCTTCGGGCTCCGACCGGTGTTCTGCGAGAAAGTCAGGA 753
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 754 ATTGGGAATTCGGTCTAGCCAGAGAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 813
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 814 CAGGGAATGTTCTCCTACGCTTCTCTCTCCCGCTTAGAGAGATGTCAGCGGAAGAGGAG 873
Qy 238 SerSerValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTFCACAGAGTTTCATCTCAGGAGGTGGCCAGACTCCCAAGGCTCCGGTACCGGAG 933
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleLysAspGlyPheAspLeuGlySer 277
Db 934 ATCATGTCCCGACACAGAAAGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
Qy 278 ValLeuAenAenAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 994 GTCCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGAGCGCTCCGCTTCAAC 1053

Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1054 CTATACGCAGTCTGCTGAGAAAGTCTCTGCTCCCTGAGTCTCTCTGATTCGTACCGTC 1113
Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1114 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGCTCTCCCGTTACCTGTAGTT 1173
Qy 338 ArgGlyLysSerGlyGluGlnArgIleHisLeuLeuGlnArgGlyIleGlyGluHis 357
Db 1174 AGAGGAATCTCCGGGAAACAAAGAAATCACTTGTCTCTGAGCGCGGATGTTGGTAGCAT 1233
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1234 CAGAAGACACAAGGGTTGCGTGCATCAAGCAACCGTGAGTGTCTGACCAAGTGCAG 1293
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 1294 GTGCGCGCGTGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGTTGGGGAG 1353
Qy 398 SerValAlaProPheAenGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1354 AGCGTCGCGCCCTTCAACCAACCGCTCACAGGCGCTGCACGCGCTTTTGTGTTTCATCAG 1413
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db 1414 CTCACCCCTCGAGCGGTGGTCCGCGCTGTCTCAATCTGGAGGAAGAGTGTCTCTGAAG 1473
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCTCGCGTATGCTGTGGAGGAGTGTGAATAGGAAGTCAGTGTGTTGACGGTGAC 1533
Qy 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTCTCTGTTTCATCATGAACATC 1593
Qy 478 LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 1594 CTTCTCCACAGACAGCCTGTGAGGAGTACTACACCTTCTTCCACCTCAGCTCTCCAGGAC 1653
Qy 498 PheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTTCTGCGCGCTTGTACTACGTGTTAGAGGGCTTGAAATCGAGCAGCTCTCTGCGCT 1713
Qy 518 LeuTyrValGlnLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTACGTTTGAGAAGACAAAGAGTCCATGAGAGCTTAAACAGGAGGCTTCCATATCCAC 1773
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 1774 TCGCTTTGGATGAAGCGTTTCTTGTGCGCTCGTGAGCGAAGACGTAAGAGGCCACTG 1833
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGTCTCTGCGGCTGCTCCCGTTCCCTGGGGGTGAAGCAGAAGCTTCTGCACTGGTTC 1893
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTGGTTCAGCAGCCTAATGCGCACCCAGGAGACACCTCGGACGCTTCCAC 1953
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAenSerPheGlnGlu 617
Db 1954 TGTCTTTTCAGACTCAAGCAAGAGTGTGTTGCGTTGGCATTTAAACAGCTTCCAGAA 2013
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGCTTCGATTAACCAAGACCTGGACTTGTATAGCATCTTCTTCTGCTCCAGCAC 2073
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCCGTATTTCGGGAAATTCGGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGATCC 2133
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677

Db 2134 GCTGAGGATGCTCTGCTGCTCTATATGAGTGGCGGATAAGACCCCTCATTTGAGGAGCAG 2193
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTCTGCTCCATGCTTGGCACCACCCACACCTGGCGCAGCTGCACCTGGGC 2253
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCGATCCCAACC 2313
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAAGATACAGACCTCATGTTTGAAGATGCACAGATTACCCCTGGTGTGCAGCACCTC 2373
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGGAGNATCGTCATGGCCACCGTNACCTAAGATCCCTCAACTGGGAGGACCCACCTG 2433
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2434 AAGGAAGAGGATGTAAGGATGGCGTGTGAAGCTTTAAACACCCCAAAATGTTTGTGAG 2493
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCCTTGAGGCTGGATGCTGTGGATTGACCATGCTCTTACCTGAAGATCTCCCAATC 2553
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTAGACCTCCCGAGCTGAAATCTCTGAGCTGGCAGGAAACAAAGTGCAGACACAG 2613
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAGTAATGCCTCTCAGTGATGCCTTGAGAGTCTCCAGTGGCCCTGCAGAGCTGATA 2673
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGGAGGACTGTGGCATCACAGCCAGCGGTGGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2734 AACCGGAGCTTGACACACTGTGCCTATCCAAACACAGCTGGGGAACGAAGGTGTAAT 2793
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2794 CTACTGTGTCATCCATGAGCTTCCCCACTGTAGTCTGCAGAGCTGATGCTGAATCAG 2853
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2854 TGCCACCTGGACACGCGCTGGCTGTGGTCTTCTGCACTTGGCTTATGGGTAACTCATGG 2913
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 2914 CTGAGCGACCTGAGCCTTAGCATGAACCTGTGAAGACAATGGCGTGAAGCTTCTGTGC 2973
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 2974 GAGGTATGAGAAACCATCTGTCACTCCAGACCTGGAGTTGGTAAAGTGTCACTC 3033
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3034 ACCGCGCGTGTGTGAGAGCTGTGCTGTGTGTCTGCACTTGGCTTATGGGTAACTCATGG 3093
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3094 CTGATCTCAACGGAATGCGCTGGGTGACGCTGGGTGTGCTGACATGTGCGAGGAGCTG 3153
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3154 AAGCAAAAGACAGTGTCTGACAGACTCGGGTTGAAGGCATGTGAGTCACTTCTGAT 3213
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037

Db 3214 TGCTGTGAGGCACTCTCCTTTGGCCCTTTCTCTGCAACCGGCATCTGACCAGTCTAAACCTG 3273
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTGCAGATTAACCTTCAGTCCCAAGGAATGATGAGCTGTGTTGCGCTTTGCTGTGCC 3333
Qy 1058 ThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
Db 3334 ACGTCTAATTACAGATAATTTGGCTGTGGAATGGCAGTACCCTGTGCAATAAGGAAG 3393
Qy 1078 LeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3394 CTGTGAGGAAGTGCAGTACTCAAGCCCGAGTCGTAATTGACGTTAGTTGCAATTCT 3453
Qy 1098 PheAspGluAspAspArg 1103
Db 3454 TTTGATGAGATGACCGG 3471
RESULT 2
AY154460
LOCUS 3885 bp mRNA linear PRI 20-FEB-2003
DEFINITION Homo sapiens NALP5 (NALP5) mRNA, complete cds.
ACCESSION AY154460
VERSION AY154460.1 GI:28436363
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NALP5: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
MEDLINE 22451042
PubMed 12563287
REFERENCE 2 (bases 1 to 3885)
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
Location/Qualifiers
1. .3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .3885
/genes="NALP5"
/genes="NALP5"
1. .3603
/notes="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains; similar
to mouse MATH protein"
/codon_start=1
/product="NALP5"
/protein_id="AA018156.1"
/db_xref="GI:28436364"
translation="MKVAGGLELGAALISASPRALVTLSTGTCTSLPKNPPLPQNL
SSQPCIMEDKSLTSSYGLQWCLYELDKREFQTKELKKDKSSSTSCISIQFETE
NANVECALLLHEYYGASLAWATSISIFENNNLRTLSSEKARDDMKRHSDEPDMEGA
QGSKEVKVIGISQAVQDSATAATKEQISQAMEQEGATAETEQEI SQAMEQEG
TAATTEQGGGTWDYKSHVMTKPAEEDVRRSFENTADWPMTLAGAFSDRWG
FRPTVVLHGKSGIGKSALARRIVLCWAQGLYQGMFVYVFLPVRMOKKSSVTE
FISREWPDAQVTEIMSPERLLFIIDGDDLLGSVLNNDTKLCKDWAQKQPTTLIR
SLURKVLPPESFLIVTVRDVGTBKKEVSVYSLVIRGSGISGQRILLERGERHQ
QTGLKRAIMNNRELLDQCPVAGSLICVALQDVGVSVAPFNQTLTGHAFAFVH
QLTPRGVVRVCLNLEERVKLRCFMAVGVWNRKSVFVGGDDLMVQGLSESLRALPH
NNLLPSDSCHEEYTFHLSQDFCAALYYVLEIEPALCPLVYEKTKRSMELKQA
GFHISLWKKRFLFGLVSDVRPLVLLGCPVLGKVKOKLLHWVSLGQPNATTPG
DTLDAFCHLFTQDKFVRLLNSFOEWLPIQNQLDLTASSFCLOHCPYLRKRVDP
KTLFPDESAECVPLMWRDKTLLIEQWEDFCMLGTHPHURKLDLGGSSILITERAM
ACBALKHPKLLSRLDCCGLTHACYLKIQTLLTSPSLSLAGNKTDOGVMLP
SDALRVSQALQKLILEDGCTATGCSLASALVSNSRLTHLCLSNNSLNGEWNLLC

MSRLPHCSLQRLMLNCHLDTPAGCFALALMGNSMLTHLSLWNPVEDNGVKLCE
VWRPESCHLOLELVKCHLTAAACSESLSVSRRHLSKSLDLTONALODGVAALCSG
LKQNSVLRLGLKACAGTSDCCALSALUSCNRHLSLNLVQNNFSPKGMKLCSAF
ACPTNSUQIIIGLWKQYVPOIRKULEEVQLLKPRVWDGSHWSFDEDDRYWKN"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3885
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
Dbs: 9 Gaps: 0

US-10-066-521-6 (1-1344) x AY154460 (1-3885)

Qy	98	IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGlu	117
Db	568	ATTTTCAAGCTATGGAACAAGAGTCCACAGCAGCAGACAGAGAACAAGAAATT	627
Qy	118	SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGluGly	137
Db	628	TCACAAGCTATGGAACAAGAGTCCACAGCAGCAGACAGAGAACAAGAGCATGGA	687
Qy	138	GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal	157
Db	688	GGTGACACATGGGACTACAGAGTCACGTGATGACCAAAATTCGCTGAGGAGGAGTGTA	747
Qy	158	ArgArgSerPheGluAunThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla	177
Db	748	CGTCGTAGTTTTGAAAAACACTGCTGCTGACTGGCCGGAATGCAAAAGTTGGCTGCT	807
Qy	178	PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly	197
Db	808	TTTGATTTCAGACCGGTGGGGCTTCGGGCTCGCCAGCGGTGTTCTGCACGGAAGTCAGGA	867
Qy	198	IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr	217
Db	868	ATTGGGAATCGGCTCTAGCCAGAAAGATCGTGCTGCTGGCGCAAGGTGGACTCTAC	927
Qy	218	GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlu	237
Db	928	CAGGGAATGTTCTCTACGTCTCTCTCTCCCGCTTACAGAGATGTCAGCGGAAGAGAG	987
Qy	238	SerSerValThrGluPheLysArgGluTrpProAspSerGlnAlaProValThrGlu	257
Db	988	AGCAGTGTTCACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTGACGGAG	1047
Qy	258	IleMetSerArgProGluArgLeuPheIleLeuAspGlyPheAspAspLeuGlySer	277
Db	1048	ATCATGTCGCCACCCGAAAGGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT	1107
Qy	278	ValLeuAunAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr	297
Db	1108	GTCTCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAGAGCAGCCTCCGTTCA	1167
Qy	298	LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal	317
Db	1168	CTCATACACGCTGCTGAGGAAGTCTGCTGCTGCTGAGTCTCTCTGATGTCACCGCTC	1227
Qy	318	ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal	337
Db	1228	AGAGACGTGGGCACNAGAAAGCTCAAGTCAGAGGTGCTGCTCTCCCGTTTACCTGTTAGTT	1287
Qy	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis	357
Db	1288	AGAGGAATCTCCGGGGAACAAGAATCCACTGCTCTTTCAGCGCGGATTTGGTGAGCAT	1347
Qy	358	GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln	377
Db	1348	CAGAAAGACACAGAGGTTCGGTGGCATCATGAACACCGTGAGCTGCTCGACAGTGCAG	1407
Qy	378	ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu	397

Db	1408	GTGCCCCCGCTGGGCTCTCTCACTCGGTGGCCCTCGAGCTGCAGGACGTGGTGGGGAG	1467
Qy	398	SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln	417
Db	1468	AGCGTGGCCCCCTTCAACCAACCGCTCACAGGCTGCACCGCGCTTTTGTGTTTCATCAG	1527
Qy	418	LeuThrProArgGlyValValArgArgCysLeuAunLeuGluGluArgValValLeuLys	437
Db	1528	CTCACCCCTCGAGCGGTGGTCCGCGCTGTCTCAATCTGAGGAAAGAGTGTCTCTGAAG	1587
Qy	438	ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp	457
Db	1588	CGCTTCTGCCGTATGGCTGTGGAGGAGGTGTGAATAGGAAGTCAGTGTTCAGCGGTGAC	1647
Qy	458	AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle	477
Db	1648	GACCTCATGCTTCAAGGACTCGGGAGTCTGAGCTCGCTCTCTGTTTCATGAAACATC	1707
Qy	478	LeuLeuProAspSerHisCysGluGlyTyrThrPhePheHisLeuSerLeuGlnAsp	497
Db	1708	CTTCTCCACAGACCCACTGTGTAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC	1767
Qy	498	PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro	517
Db	1768	TTCTGTGCCCTTGTACTACGTGTAGAGGCTGTGAATTCAGACGAGTCTCTGCCCT	1827
Qy	518	LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis	537
Db	1828	CTGTAGCTTCAGAAAGACAAAGAGTCCATGGAGCTTAAACAGGAGCTTCCATATCCAC	1887
Qy	538	SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu	557
Db	1888	TCGCTTTTGGATGAAGCGTTTCTTGTGTGGCTCGTGAGCGAAGACGCTAAGAGCGCATG	1947
Qy	558	GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal	577
Db	1948	GAGTCTCTGTGGCTGTCCGCTTCCCTGGGGGTGAGCAGAGCTTCTGCACTGGGTC	2007
Qy	578	SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis	597
Db	2008	TCCTGTGGTTCAGCAGCCTAATGTCACCCAGGAGACACCTCGGACGCTTCCAC	2067
Qy	598	CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAunSerPheGlnGlu	617
Db	2068	TGCTTTTTCAGACTCAAGCAAGAGTGTGTGCTTGGCATTTAAACAGCTTCCAGAA	2127
Qy	618	ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis	637
Db	2128	GTGTGGCTTCCGATTAACAGAACCTGGACTTGTATAGCATCTTCTTCTGCTTCCAGAC	2187
Qy	638	CysProTyrLeuArgLysIleArgValLysGlyIlePheProArgAspGluSer	657
Db	2188	TGTCGCTATTTCGGGAAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGATGAGTCC	2247
Qy	658	AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln	677
Db	2248	GCTGAGGACATGTCCTGTGGTCCCTCTATGATGCGGATAAGACCTTCATTGAGGAGCAG	2307
Qy	678	TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly	697
Db	2308	TGGGAAGATTCTGCTCCATGCTTGGCACCCACCCACACTCGCGGACGTGGACCTGGGC	2367
Qy	698	SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr	717
Db	2368	AGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACC	2427
Qy	718	CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu	737
Db	2428	TGCAAGATACAGACCTTGTATTTAGAAATGCAAGATTACCCCTGGTGTGACGACCTC	2487
Qy	738	TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu	757

```
Db 2488 TGGAGAAATCGTCATGGCCCAACCGTAACCTAAGATACCTCAACTTTGGGAGCACCACCACTG 2547
Qy LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2548 AAGGAGAGGATGTAGAGATGGCTGTGAACCTTAAACACCCCAAAATGTTGTTGGAG 2607
Qy SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2608 TCTTTAGGCTGGATGCTGTGGATGACCCATGCCGTGTACCTGAAGATCTCCCAATC 2667
Qy LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2668 CTTAGACCTCCCCAGCTGAATCTCTGAGCCTGGCAGAAACAAGGTGACAGACAG 2727
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2728 GGAGTAATGCTCTCAGTGATGCTTGGAGTCTCCAGTGGCCCTGCAGAGCTGATA 2787
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2788 CTGGAGGACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2847
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2848 AACCGGAGCTTGACACACCTGTGCTATCCAAACAACAGCCTGGGGAAACGAAGGTGTAAT 2907
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2908 CTACTGTGTGCATCATGAGGCTTCCCCACCTGTAGTCTGCAGAGCCTGTAGTCTGAATCAG 2967
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2968 TGCCACCTGGACACGGCTGGCTGTGTTCTTTCACCTTGGCTTATGGGTAACTCATGG 3027
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 3028 CTGACGCACCTGAGCCTTAGCATGAACCTCTGGAAGACAATGGCGTGAAGCTTCTGTGC 3087
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluValLysCysHisLeu 957
Db 3088 GAGGTGATGAGAGAACCATCTGTGTCATCTCCAGACCTTGGAGTTGGTAAAGTGTATCTC 3147
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3148 ACCGCCGGCTGTGTGAGAGCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 3207
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3208 CTGGATCTTCACGCAAAATGCTGCGTGCAGCGTGGGGTTGCTGCACCTGTGCGAGGACTG 3267
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3268 AAGCAAAAGAACAGTGTTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGAT 3327
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 3328 TGCTGTGAGGCACCTCTCTTGCCCTTTCCTGCAACCGCATCTGACCACTGTAAACCTG 3387
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3388 GTGCAGAAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCCCTTTGCTGCTCCC 3447
Qy 1058 ThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
Db 3448 ACGTCTAACTTACAGATAAATGGGCTGTGGAAATGGCAGTAGTACCCTGTGCAAAATGAAGAG 3507
Qy 1078 LeuLeuGluGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3508 CTGCTGGAGGAAGTGCACACTACTCAAGCCCGCAGTCTGTAATTGACGTAGTTGCACTTCT 3567
Qy 1098 PheAspGluAspArg 1103
Db 3568 TTTGATGAAGATGACCGG 3585
```

```
RESULT 3
AX704821
LOCUS AX704821 3926 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 1 from Patent EP1285964.
ACCESSION AX704821
VERSION AX704821.1 GI:29561487
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Weiss,B., Lessl,M., Peters-Kottig,M. and Beckmann,G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964-A 1 26-FEB-2003;
SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES
Location/Qualifiers
source 1..3926
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3926
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0
US-10-066-521-6 (1-1344) x AX704821 (1-3926)
Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIle 117
Db 454 ATTTCAAGACTATGAAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGAAATT 513
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
Db 514 TCACAGACTATGAAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGGACATGGA 573
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 574 GGTGACACATGGGACTACAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGGATGTA 633
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTGTAAACACTGCTGCTGACTGCGCCGAAATGCAAACTGGCTGGTGTCT 693
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 694 TTTGATTCACAGCGGTGGGCTTCCGGCTTCGACCGTGGTCTGACACGGAAGTTCAGGA 753
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 754 ATTTGGGAATCGCTCTAGCCACAAGGATCGTCTGTGCTGGCGCAAGGTGGACTCTAC 813
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 814 CAGGGAATGTTCTCTACGTTCTTCTCTCCCTCCCGTTAGAGAGATGCAAGCAAGAGAG 873
Qy 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTACAGAGTTCACTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGGTGACGGAG 933
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 934 ATCATGTCCCGACCAAGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
Qy 278 ValLeuAsnAsnAspThrLysLeuLysLysAspTrpAlaGluLysGlnProPheThr 297
Db 994 GTCTCTCAACAATGACACAAAGCTCTGCAAGAAGCTGGGCTGAGAAGCAGCCTCGCTTCA 1053
```

QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 1054 CTCATACGAGTCTGCTGAGGAGGTCTGCTCCCTGAGTCTCTCCTGATCGTCAACGTC 1113
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
DB 1114 AGAGACGTGGGCACAGAGAAAGCTCAAGTCAGAGGTCTGTCTCCCGGTACCTGTAGTT 1173
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
DB 1174 AGAGAAATCTCCGGGAACAAGAAATCCATGTCTCTTGAGCGGGAGATTGGTGAGCAT 1233
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
DB 1234 CAGAAGACACAGAGGTTCGGTGGCATCATGAACACCGTGAGCTGCTGCACAGTGCAG 1293
QY 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
DB 1294 GTGCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGTGGGGAG 1353
QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
DB 1354 AGCGTCGCCCGCTTCAACCAACCGCTCACAGCGCTGCACGCGCTTTTGTCTTCATCAG 1413
QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
DB 1414 CTCACCCCTCAGAGCGTGGTTCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTCGAAG 1473
QY 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB 1474 CGCTTCTCCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCAGCGGTGAC 1533
QY 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
DB 1534 GACCTCATGTGTCAAGGACTCGGGAGTCTGAGTCCGTCTCTGTTTCATGAACATC 1593
QY 478 LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497
DB 1594 CTTTCTCCAGACAGCCACTGTGAGGAGTACTACCTTCTTCCACCTCAGTCTCCAGGAC 1653
QY 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB 1654 TTCGTGCGCGCTTGTACTACGTGTAGAGGGCTGGAAATCGAGGACGCTCTCTGCCCT 1713
QY 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB 1714 CTGTACGTTGAGAAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGGCTTCCATATCCAC 1773
QY 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
DB 1774 TCGCTTTGGATGAAGCGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAGAGGCCACTG 1833
QY 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
DB 1834 GAGGTCTGCTGGGCTGTCCTGCTCCCTGGGGTGAAGCAGAGCTTCTGCATGGGTC 1893
QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 1894 TCTCTGTTGGTTCAGACAGCTTAATGCCACCAACCCAGGAGACACCTTGGAGCGCTTCCAC 1953
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 1954 TGTCTTTTCAGACTCAAGACAAAGAGTTTGTTCGTGGCATTAACAGCTTCCAAAGAA 2013
QY 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 2014 GTGTGGCTTCCGATTAAACCAACCTGGACTTGTAGCATCTTCTCTGCTCCAGCAC 2073
QY 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 2074 TGTCCGTAATTTGCGGAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC 2133

QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 2134 GCTGAGGCATGTCTGTGTCCTCTATGATGCGGGATGAACCCCTCATTTGAGGACGAG 2193
QY 678 TrpGluAspPheCysSerMetLeuLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 2194 TGGGAAGATTCTGCTCCATGCTTGGCACCCACCCACCTGCGGACGTGGACCTGGGC 2253
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 2254 AGAGCATCTCTGACAGAGCGGCCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACC 2313
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 2314 TGCAGATACAGACCTGTATGTTTAGAATGCAAGATTACCCCTGCTGGTGTGACACCTC 2373
QY 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
DB 2374 TGGAGAATCGTTCATGCGCAACCGTAACTTAAGATCCCTCAACTTGGGAGCACCCACCTG 2433
QY 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
DB 2434 AAGGAAGAGATGTAAGGATGGCGTGTGAAGCCTTAAAAACACCAAAATGTTTGTGGAG 2493
QY 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
DB 2494 TCITTTGAGGCTGAGTTGCTGTGATTTGACCATGCTGTGTACCTGAAGATCTCCCAATC 2553
QY 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
DB 2554 CTTTACGACCTCCCGGAGCTGAAATCTCTGAGCCTGGCAGGAACAAGGTGACAGACCAG 2613
QY 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
DB 2614 GGAATTAATGCTCTCAGTCATGCTTTCAGAGTCTCCAGTGCCTGCGCCCTGCAGAAAGCTGATA 2673
QY 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
DB 2674 CTGGAGACTGTGGCATCACAGCACGGGTTCAGAGTCTGGCCCTCAGCCCTCGTCAGC 2733
QY 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
DB 2734 AACCGGAGCTTGACACACCTGCTGCTATCCAAACAACAGCCTGGGGAACAAGGTGTAAT 2793
QY 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
DB 2794 CTACTGTGTCATGCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
QY 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
DB 2854 TGCACCTGGACACGCTGGCTGTGTTTTCTTGTGCACTTGGCTTATGGGTAACTCATGG 2913
QY 918 LeuThrHisLeuSerSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
DB 2914 CTCACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAAATGGCGTGAAGCTTCTGTGC 2973
QY 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeu 957
DB 2974 GAGGTTCATGAGAGAACATCTTGTCTCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
QY 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
DB 3034 ACCCGCGGTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGACACACCTGAAGAGC 3093
QY 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
DB 3094 CTGGATCTCAAGCACAAATGCCCTGGGTGACGGTGGGGTGTGCTGCACTGTGCGAGGACCTG 3153
QY 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
DB 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTCGAAGGCATGTGAGCTTCTGAT 3213
QY 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037

```
Db 3214 TCCTGTGAGGCACTCTCTCTTGGCCCTTTCTTCCGCAACCGGCATCTGACCAAGTCTAAACCTG 3273
Qy ValGlnAsnAsnPhSerProlysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTGCAGATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCCCTTTGCTGTCTCC 3333
Qy 1058 ThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGlnTrpProValGlnIleArgLys 1077
Db 3334 ACGTCTAACTTACAGATAATTGGGCTGTGGAAATGGCAGTACCCTGTGCAAAATAAGGAAG 3393
Qy 1078 LeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3394 CTGCTGGAGGAAGTCAGCTACTCAAGCCCGAGTCGTAATTGACGGTAGTTGGCATTTCT 3453
Qy 1098 PheAspGluAspAspArg 1103
Db 3454 TTTGATGAAGATGACCGG 3471

RESULT 4
AX704823 3030 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 3 from Patent EP1285964.
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Weiss,B., Lessl,M., Peters-Kottig,M. and Beckmann,G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964-A 3 26-FEB-2003;
SCHERING AKTIENGESSELLSCHAFT (DE)
FEATURES
source
1..3830
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3830
Score: 988.00 Matches: 988
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.51% Indels: 0
Db: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX704823 (1-3830)
Qy 116 GlutSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGly 135
Db 451 GAAATTTTCAAGCTATGTGAACAGAGAGTCCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAG 510
Qy 136 HisGlyGlyAspThrTrpAspTrpLysSerHisValMetThrLysPheAlaGluGluGlu 155
Db 511 CATGGAGGTGACATGGGACTACAGAGTACAGTACGATGATGACCAATTCGCTGAGAGAGAG 570
Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 571 GATGTACCTCGTAGTTTGAAGAACACTGCTGCTGACTGGCCGGAATGCAAAACGTTGGCT 630
Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 631 GGTGCTTTTGATTCAGACCGGTGGGCTTCGGGCTCGCACGGTGGTTCGTCACGGAAG 690
Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgIleValLeuLysCysTrpAlaGlnGlyGly 215
Db 691 TCAGGAATTTGGGAATCGGCTCTAGCCAGAGAGATCGTGTGCTGGGCGCAGAGGTGGA 750
Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
```

```
Db 751 CTCTACAGGGAATGTTCTCTACGTCTTCTTCCCTCCGTTAGAGAGATGACGCGAAG 810
Qy LysGluSerSerValThrGluPheIleSerArgGlnTrpProAspSerGlnAlaProVal 255
Db 811 AAGGAGAGCAGTGTTCACAGAGTTTCATCTCCAGGAGGTGGCCAGACTCCCAAGGCTCCGGTG 870
Qy 256 ThrGluIleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeu 275
Db 871 ACGGAGATCATGTCCCGACCCAGAAAGGCTGTGTTCATCATTTGACGGTTCGATGACCTG 930
Qy 276 GlySerValLeuAsnAsnAspThrLysLeuLysCysLysAspTrpAlaGluLysGlnProPro 295
Db 931 GGCTCTGTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGGGCTGAGAAAGCAGCTCCG 990
Qy 296 PheThrLeuIleArgSerLeuLeuArgLysValLeuLeuLeuProGluSerPheLeuIleVal 315
Db 991 TTTACCTCTCATACGAGTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCTTCTCTGATCGTC 1050
Qy 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
Db 1051 ACCGTACAGACAGTGGGSCACAGAGAGCTCAAGTTCAGAGGTCGTGTCTCCCGTTACCTG 1110
Qy 336 LeuValArgGlyLysSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGly 355
Db 1111 TTAGTTAGAGGAATCTCCGGGGAACAAAGAAATCCACTTGTCTCTTGGAGCGCGGATTTGGT 1170
Qy 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
Db 1171 GAGCATCAGAAAGACACAAAGGTTGCGTGCATCATGAACAAACCGTGAAGTGTGCTGACGAC 1230
Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db 1231 TGCCAGGTGCGCGCGTGGGCTCTCTCATCTGCTGGCCCTGCAGCTGCAGGAGCTGGTG 1290
Qy 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
Db 1291 GGGGAGAGCGTCGCCCTTCAACCAACCGCTCACAGGCTGCACGCGCTTTTGTGTTT 1350
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValVal 435
Db 1351 CATCAGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGGAGGAAAGAGTTGTC 1410
Qy 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
Db 1411 CTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGAAATAGGAAGTCAAGTGTTCAC 1470
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
Db 1471 GGTGACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTCACATG 1530
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeu 495
Db 1531 AACATCTCTTCCACAGACGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1590
Qy 496 GlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeu 515
Db 1591 CAGGACTTCTGTCGCCCTTGTACTACGTGTGTAGAGGGCTCGGAATCGAGCCAGCTCTC 1650
Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db 1651 TGCCCTCTGTACGTTGAGAAGACAAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCAT 1710
Qy 536 IleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
Db 1711 ATCCATCGCTTTGGATGAAGCGTTTCTTGTGTTGGGCTCGTGGAGCGAAGACGTAAGGAGG 1770
Qy 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
Db 1771 CCACTGGAGGTCCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGACTTCTGTCAC 1830
Qy 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
```

Db 1831 TGGGTCTCTCTGTTGGGTGACGAGCCTAATGCCACCACCCAGGAGACACCCCTGGACGCC 1890

Qy 596 PheHisCysLeuPheGluThrClnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615

Db 1891 TTCCACTGTCTTTTCGAGACTCAGACAAAGAGTTTGTTCCTTGGCATTTAAACAGCTTC 1950

Qy 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635

Db 1951 CAAGAGTGTGGCTTCCGATTAAACAGAACTGGACTTGTATAGCATCTTCTCTTCGCTC 2010

Qy 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655

Db 2011 CAGCACTCTCCGTATTTTCGGAAATTCGGTGGATGTCAAGGGATCTTCCCAAGAGAT 2070

Qy 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675

Db 2071 GAGTCCGCTGAGGCATGTCCTGTGGTCCCTATGTGATGCGGATAAGACCCCTCATTTGAG 2130

Qy 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuAlaGlnLeuAsp 695

Db 2131 GAGCAGTGGAGAGATTCTGTCTCATGCTTGGCACCCACCCACACCTGGCGCAGCTGGAC 2190

Qy 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715

Db 2191 CTGGGCAGCAGCATCTGACAGAGCGGGCCATGAAGACCCCTGTGTGCCAAGCTGAGGCAT 2250

Qy 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735

Db 2251 CCCACCTGCAAGATACAGACCTCATGTTTGAATAATGCACAGATTACCCCTGGTGTGCAG 2310

Qy 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThr 755

Db 2311 CACCTCTGGAGAGATCGTCATCGGCCAACCGTAACCTAGATCCCTCAACTTGGGAGGACCC 2370

Qy 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775

Db 2371 CACCTGAAGGAAGAGATGTAAGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTTG 2430

Qy 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795

Db 2431 TTGAGTCTTTTGAGCGTGATGTGTGTGATGTGACCCATGCCATGCTTACCTGGAAGATCTCC 2490

Qy 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThr 815

Db 2491 CAATCTCTAGACCTCCCCAGAGCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACA 2550

Qy 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835

Db 2551 GACCAGGGAGTAAATGCCCTCTCAGTGATGCTTGAAGATCTCCAGTGGCCCTGCAGAAG 2610

Qy 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855

Db 2611 CTGATACTGGAGGACTGTGGGCATCACAGCCACGGGTTGCCAGAGTCTGGCCCTCAGCCCTC 2670

Qy 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875

Db 2671 GTCAGCAACCGAGCTTGACACACTGTGCCCTATCCAAACACACCTTGGGGAAACAGGT 2730

Qy 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895

Db 2731 GTAAATCTACTGTGTGATCCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATGCTG 2790

Qy 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheIleuAlaLeuAlaLeuMetGlyAsn 915

Db 2791 AATCAGTGCACCTGGACACGGCTGGCTGTGGTGTCTTTCACACTTGGCTTATGGGTAAAC 2850

Qy 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935

Db 2851 TCATGGCTGAGCAGCCTGAGCCTTAGCATGAAACCTCTGTGGAGACATGGCTGGAACCTT 2910

Qy 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955

Db 2911 CTGTGCGAGGTTCATGAGAGAACCATTCTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGT 2970

Qy 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975

Db 2971 CATCTCACCCCGCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGACACACCTG 3030

Qy 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGlu 995

Db 3031 AAGAGCTGGATCTCAGGCACATGCTGGGTGACGGTGGGTGCTGCATCTGTGGCAG 3090

Qy 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015

Db 3091 GGACTGAAGCAAAAGAACAGTGTCTTGACGAGACTCGGGTTGAAGGCATGTGACTGACT 3150

Qy 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035

Db 3151 TCTGATGTCTGAGGCATCTCTCTTGGCCCTTCTGCAACCGGCATCTGACGACGCTA 3210

Qy 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055

Db 3211 AACCTGGTGAGATTAATCTTCAGTCCCAAGGAATGATGAGCTGTGTTCGGCTTTGCC 3270

Qy 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075

Db 3271 TGTCCACAGTCTAATCTACAGATAATTTGGCTGTGAAATGGCAGTACCTGTGCAATA 3330

Qy 1076 ArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrp 1095

Db 3331 AGGAAGCTGTGTGAGGAAGTGCAGCTACTCAAGCCCGAGTCGTAATTACGGTAGTTGG 3390

Qy 1096 HisSerPheAspGluAspArg 1103

Db 3391 CATCTTTTGATGAAGATGACCGG 3414

RESULT 5

AX459881 LOCUS 5859 bp DNA linear PAT 30-NOV-2002
SEQUENCE 24 from Patent WO0240668.

AX459881 ACCESSION

AX459881.1 GI:21725649

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Tschoopp,J. and Martinon,F.

Proteins and dna sequences underlying these proteins used for

treating inflammations

Patent: WO 0240668-A 24 23-MAY-2002;

Apotech Research and Development Ltd. (CH)

JOURNAL

FEATURES

source

1..5859
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP8/Py12---Py8.dna"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 5859
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX459881 (1-5859)

Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117

Db 2389 ATTTCAGAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGAGAACAAGAAATT 2448

Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137

Db 2449 TCACAAAGCTATGGAAACAAGAGGTGCCACAGCAGACAGAGAACAAAGGACATGGGA 2508
Qy GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspVal 157
Db 2509 GGTGACACATGGGACTCAACAGATCACGTGATGACCAAAATTCGCTGAGGAGGAGATGTA 2568
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 2569 CGTCGTAGTTTTGAAAAACACTGCTGCTGACTGGCGGCGGAAATGCAAAACGTTGGCTGGTCT 2628
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 2629 TTGTGANTCAGACCGGTGGGGCTTCGGCTCGCACGGTGGTTCGCAACGGAAGTCAGGA 2688
Qy 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 2689 ATTGGGAATCGGCTCTAGCCAGCAAGGATCGTGTGTGCTGGGCGCAAGTGGACTCTAC 2748
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlu 237
Db 2749 CAGGGAATGTTCTCTACGTCTTCTTCTCCCGTTAGAGAGATGTCAGCGCAAGAGAG 2808
Qy 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 2809 AGCAGTGTACAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCCAAGGCTCCGGTGACGGAG 2868
Qy 258 IleMetSerArgProGluArgLeuLeuPheIleLeuAspGlyPheAspAspLeuGlySer 277
Db 2869 ATCATGTCCCGACCCAGAAAGCTGTTGTTCAITGACGGTTTCGATGACTGGGCTCT 2928
Qy 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 2929 GTCCCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAGAGCAGCCTCCGTTCAAC 2988
Qy 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 2989 CTCATACCCAGTCTGCTCAGGAAGTCTCTGCTCCCTGAGTCTCTTCATCGTCAACCGTC 3048
Qy 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 3049 AGAGACGTGGGCACAGAGAACTCAAGTCAGAGGTCTGTCCTCCCGTTACTCTGTAGTT 3108
Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 3109 AGAGGAATCTCCGGGAAACAAGAAATCCACTTGCTCTTGTAGCGCGGGATTGGTGAGCAT 3168
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 3169 CAGAAGACACAAGGTTTCGTCGCGCATCATGAACAACCGTGAGCTGCTCGACCACTGGCCAG 3228
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 3229 GTGCCCGCGTGGGCTCTCATCTGCGTGGCCCTGACAGCTGACAGGAGTGGTGGGGGAG 3288
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 3289 AGCGTCGCCCCCTTCAACCAACCGCTCACAGCCCTGCACGCGCTTTGTGTTCATCAG 3348
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db 3349 CTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCGAAG 3408
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 3409 CGCTTCTCCCGTATGGCTGTGGAGGAGTGGGAATAGGAAGTCAGTGTTTTGACGGTGAC 3468
Qy 458 AspLeuMetValGlnGlyLeuGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 3469 GACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTGCTGTTCATCATGACATC 3528
Qy 478 LeuLeuProAspSerHisCysGluGluTyrTrpPhePheHisLeuSerLeuGlnAsp 497

Db 3529 CTTCTCCACAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 3588
Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluLeuProAlaLeuCysPro 517
Db 3589 TTCTGTGCCGCTTGTTACTACTAGTGTAGAGGGCTTGAATCGAGCCAGCTCTCTGCCCT 3648
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 3649 CTGTACGTTTGAAGAACAAGAGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCAC 3708
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 3709 TCGCTTTGGATGAAGCGCTTCTTGTGTGGCTCGTGAGCGAAGACGTAAGAGGCCACTG 3768
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 3769 GAGTCTCTGTGGCTGTCCCGTTCCTCGGGGTGAAGCAGAAGCTTCTGCCTGGGTC 3828
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 3829 TCTCTGTGGTTCAGCAGCCTAATGCCACCCAGGAGACACCTTGGACGCTTCCAC 3888
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 3889 TGTCTTTTCGAGACTCAAGACAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAAGAA 3948
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 3949 GTGTGGCTTCGATTAACAGAACCTTGGACTTGTATAGCATCTTCTTCTGCCTCCAGC 4008
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 4009 TGTCCGTAATTTGGGAAAATTCGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCC 4068
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrIleGluGlnGln 677
Db 4069 GCTGAGGACATGCTCTGTGTCCCTCTATGATGCGGATGAAGACCTTCATTGAGGAGCAG 4128
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 4129 TGGGAAGATTTCTGCTCCATGCTTGGCACCCACCCACACCTGCGGACAGCTGGACCTGGGC 4188
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 4189 AGCAGCATCTGCAGACGCGGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCAC 4248
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 4249 TGCAAGATACAGACCCCTGATGTTTAGAAATGCAAGATTACCCCTGGTGTGCAGCACCTC 4308
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu 757
Db 4309 TGAGAAATCTCATGGCAACCGCTAACCTAAGATCCCTCACTTGGAGGAGCACCACCTG 4368
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 4369 AAGGAAGAGATGTAGGATGGCGTGTGAAGCTTAAACACCAAAATGTTTGTGGAG 4428
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 4429 TCTTTGAGGCTGGATTTGCTGTGGATTTGACCCATGCTTACCTGAAGATCTCCAAATC 4488
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 4489 CTTACGACCTTCCCGCAGCCTGAAATCTCTGAGGCTGSCAGAAACAAGGTGACAGACAG 4548
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 4549 GGAGTAATGCTCTCTAGTGATGCTTGGAGAGTCTCCAGTGGCCCTGCGAGAGCTGATA 4608
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 4609 CTGAGGACTGTGGCATCACAGCCACGGGTGCCAGAGTGTGGCCTCAGCCCTCGTCAGC 4668

Qy	858	AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn	877		
Db	4669	AACCGAGCTTGACACACCTGTGCTATCCAAACACGCTGGCGGAACGAAGGTGTAAT	4728		
Qy	878	LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln	897		
Db	4729	CTACTGTGTGATCCATGAGGCTTCCCACCTGTAGTGTGCAGAGGCTGATGCTGAATCAG	4788		
Qy	898	CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp	917		
Db	4789	TGCCACCTGGACACGGCTGGCTGTGGTTTCTTGCATTTGGCTTATGGGTAACTCAGG	4848		
Qy	918	LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys	937		
Db	4849	CTGACGCACCTGAGCCTTAGCATGAACCCCTGTGGAACACAATGGCGTGAAGCTTCTGTGC	4908		
Qy	938	GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu	957		
Db	4909	GAGGTCAATGAGAGAACCATCTTGTTCATCTCCAGACCTGGAGTTGGTAAAGTGTTCATCTC	4968		
Qy	958	ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer	977		
Db	4969	ACCGCCGGTGTCTGAGAGTCTCTCTGTGTGATCTCGAGGAGCACACCTGAAGAGC	5028		
Qy	978	LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu	997		
Db	5029	CTGGATCTCAGGACAATGCCCTGGGTGACGGTGGGTTGCTGCACCTGTGCGAGGACTG	5088		
Qy	998	LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp	1017		
Db	5089	AAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGACTGACTTCTGAT	5148		
Qy	1018	CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu	1037		
Db	5149	TGCTGTGAGGACACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTAAACCTG	5208		
Qy	1038	ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro	1057		
Db	5209	GTGCAGATAAATTCAGTCCCAAGGAATGATGAAGCTGTGTCGGCTTTGCTGTCTCCC	5268		
Qy	1058	ThrSerAsnLeuGlnIleGly 1065			
Db	5269	ACGTCTAACTTACAGATAATTGGC 5292			
RESULT 6					
AX459873	AX459873	6939 bp	DNA	linear	PAT 30-NOV-2002
LOCUS	Sequence 16 from Patent WO0240668.				
DEFINITION	AX459873				
ACCESSION	AX459873.1	GI:21725645			
VERSION					
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1				
JOURNAL	Tschopp, J. and Martinon, F.				
FEATURES	Proteins and dna sequences underlying these proteins used for				
source	treating inflammations				
	Patent: WO 0240668-A 16 23-MAY-2002;				
	Apotech Research and Development Ltd. (CH)				
	Location/Qualifiers				
	1. .6939				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	/note="NALP5/Py8.dna /GENSCAN_predicted_CDS_1/ 6939_bp"				
ORIGIN					
Alignment Scores:					
Pred. No.:	0	Length:	6939		
Score:	968.00	Matches:	968		

Db 4429 CTCACCCCTCAGGCGGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCTGAAG 4488
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 4489 CGCTTCTCCGCTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAAGTGTGGAGCGGTGAC 4548
Qy 458 AspLeuMetValGlnGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 4549 GACCTCATGGTTCGAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTCTTTCACATGAACATC 4608
Qy 478 LeuLeuProAspSerHisCysGluGluTyrTrpPhePheHisLeuSerLeuGlnAsp 497
Db 4609 CTTCTCCAGACAGCCACTGTGAGAGGACTTACACCTTCCTTCCACCTCAGTCTCCAGGAC 4668
Qy 498 PheCysAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 4669 TTCTGTCCGCTGTACTACTGTGTAGAGGCTTGGAAATCGAGCCAGCTCTCTGCCCT 4728
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 4729 CTGTACGTTGAGAAACAAAGAGTCCATGAGCTTAAACAGGCAGGCTTCCATATCCAC 4788
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 4789 TCGCTTTGGATGAAGCGTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAGAGGAGCCACTG 4848
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 4849 GAGTCTCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTC 4908
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 4909 TCTCTGTGGTTCAGCAGCCTAATGCCACCAACCCAGGAGACACCTTGGAGCCTTCCAC 4968
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 4969 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTCTGGCTTGGCATTTAAACACAGCTTCCAAGAA 5028
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 5029 GTGTGGCTTCCGATTAAACCAAGACTGGACTGTGATAGCATCTTCTTCTGGCTCCAGCAC 5088
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 5089 TGTCCGATTTCGGGAAATTCGGGTGGATGTCAAAAGGAGTCTTCCCAAGAGATGAGTCC 5148
Qy 658 AlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
Db 5149 GCTGAGGATGCTCTGTGTGTCCTCTATGGATGGGGATAGACCCCTCAATTGAGGAGCAG 5208
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 5209 TGGGAAGATTTCGTCTCATGCTTGGCACCCACCCACACCTTGGCGAGCTGGACCTGGGC 5268
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 5269 AGCAGCATCTGACAGAGCGGCCCATAGACCTGTGTGCCAAGCTGAGGCATCCACCC 5328
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 5329 TGCAAGATACAGACCTGTATTTAGAAATGCACAGATTACCCCTGGTGTGTCAGCACCTC 5388
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThrHisLeu 757
Db 5389 TGGAGAATTCGTATGGCCAAACCGTAACCTAAGATTCCTCAACTTGGGAGGCCACCTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAGAGGATGTAGAGTGGCTGTGAACCTTAAACACCCAAATATGTTGTGGAG 5508
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATTGCTGTGGATTGACCCCATGCTGTACCTGAAGATCTCCCAATC 5568

Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTACGACCTCCCCAGCCTGAATCTCTCAGCCTGGCAGGAAACAGGTGACAGACCAG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAATTAATGCTCTCAGTGTATGCTTGGAGAGTCTCCAGTGGCGCTGCAAGACTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGAGGACTGTGGCATCACAGCCACGGTTCAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACCTGTGCTTATCCACACAGCCTGGGGAACGAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCACCTGGACACGGCTGGCTGTGTTTCTTTCACCTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGACCTGAGCCTTAGCATGACCTGTGGAGACAACTGGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGTTCATGAGAAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCACTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCCGCTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGGATCTCAGGACAAATGCCCTGGGTGACGGTGGGGTTGCTGCATGTGCGAGGAGTCT 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAAACAGTGTCTCAGCAGACTCGGGTTGAAGGCATGTGAGCTTCTGTAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGACACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTTAAACCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGAGATTAACCTTCAGTCCCAAGGAATGATGAGCTGTGTTGGCCTTGTCTGTCTTCCC 6348
Qy 1058 ThrSerAsnLeuGlnIleLeuGly 1065
Db 6349 ACCTTAACCTTACAGATAATTGGC 6372

RESULT 7

AX459891
LOCUS AX459891 6939 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 34 from Patent WO0240668.
ACCESSION AX459891
VERSION AX459891.1 GI:21725654

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Tschopp, J. and Martinon, F.

Proteins and dna sequences underlying these proteins used for

treating inflammations

Patent: WO 0240668-A 34 23-MAY-2002;

JOURNAL

FEATURES Apotech Research and Development Ltd. (CH)

source Location/Qualifiers
1..6939
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP13/Py17.cdna"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX459891 (1-6939)

QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluLeu 117
DB 3469 ATTTCACAAGCTATGGAAACAGAAGGTGCCACAGCAGACAGACAGAAACCAAGAAATT 3528
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
DB 3529 TCACAGCTATGGAAACAGAAGGTGCCACAGCAGACAGACAGAAACCAAGGACATGGA 3588
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
DB 3589 GGTGACACATGGGACTACAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGATGTA 3648
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
DB 3649 CGTCGTAGTTTGAACACACTGCTGCTGACTGCGCGGAAATGCAAAACGTTGGCTTGGTGC 3708
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
DB 3709 TTTGATTCAGACCGGTGGGGCTTCGGCTCGCAGCGTGGTTCGACGAAAGTCAAGA 3768
QY 198 IleGlyLysSerAlaLeuAlaArgAlaValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
DB 3769 ATTTGGAAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 3828
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
DB 3829 CAGGGAATGTTCTCCTAGCTCTTCTTCTCCCGTTAGAGAGATGCGACGCGAAGAAGGAG 3888
QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
DB 3889 AGCAGTGTACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTGACGGAG 3948
QY 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
DB 3949 ATCATGTCCCGACACAGAAAGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 4008
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
DB 4009 GTCCCTCAACAATGATACAAAGCTCTGCAAAAGACTGGGCTGAGAAAGACGCTCCGTTCA 4068
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 4069 CTCATAGCAGTCTGCTGAGGAAGTCTGTCTCTGATGCTTCTCTGATCGTCAACGCTC 4128
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
DB 4129 AGACAGCTGGGCAACAGAAAGCTCAAGTCAAGGTGCTGTGCTCTCCCGTACCTGTTAGTT 4188
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
DB 4189 AGAGGAATCTCCGGGGAAACAAAGAAATCCACTGTCTCTGAGCGCGGGATTGGTGACAT 4248
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuLeuCysGln 377

DB 4249 CAGAAGACACAAGGGTTGGTGGCATCATGAACAACCGTGAGCTGCTCGACACAGTGCAG 4308
QY 378 ValProAlaValGlySerLeuIleCysValAlaAlaLeuGlnLeuInAspValValGlyGlu 397
DB 4309 GTCCCGCCCGTGGGCTCTCTCATCTCGTGGCCCTGACAGTGCAGGACGCTGGGGGAG 4368
QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
DB 4369 AGCGTGCCCGCTTCAACCAACGCTCACAGGCTGCACGCCGCTTTTGTTTCATCAG 4428
QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
DB 4429 CTCACCCCTCGAGCGCTGGTCCGCGCTGTCTCAATCTGGAGGAAAGAGTTGCTCGAAG 4488
QY 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
DB 4489 CGCTTCTGCGCTATGGCTGTGGAGGGAGTGTGAATAGGAAGTCAGTGTGTGACGGTGC 4548
QY 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
DB 4549 GACCTCATGTTTCAAGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTCATATGAACATC 4608
QY 478 LeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAsp 497
DB 4609 CTTCTCCAGACAGCACCTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 4668
QY 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
DB 4669 TTTCTGTGCGGCTTGTACTACGTGTAGAGGGCTTGAAATCGAGCCAGCTCTCTGCGCT 4728
QY 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
DB 4729 CTGTAGCTTGAGAAACAAAGAGGTCCATGAGCTTAAACAGGACGCTTCCATATCCAC 4788
QY 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
DB 4789 TCGCTTTGGATGAAGCGTTCTTGTGTGGCTCGTGCAGGAGACCGTAAAGAGCCACATG 4848
QY 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisIleTrpVal 577
DB 4849 GAGGTCTGTGCGGCTGTCCCGTTCCCTGGGGGTGAAAGCAGAAAGCTTCTGCACCTGGTC 4908
QY 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
DB 4909 TCTCTGTGGTTCAGACGCTAATGCCACACCCAGGAGACACCTTGGACGCTTCCAC 4968
QY 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
DB 4969 TGTCTTTTCAGACTCAAGCAAAAGAGTTGTTCGCTTGGCATTAACAGCTTCCAAGAA 5028
QY 618 ValTrpLeuProIleAsnGlnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
DB 5029 GTGTGGCTCCGATTAAACAGAACCTGGACTTGATAGCATCTTCTCTTCTGCTCCAGCAC 5088
QY 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
DB 5089 TGTCCGTATTTCGGAAAAATTCGGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGATGCC 5148
QY 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
DB 5149 GCTGAGGCATGCTCTGTGGTCCCTCTATGGATGCGGGATAAGACCTCATTTGAGGAGCAG 5208
QY 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
DB 5209 TGGGAAGATTCTGCTCCATGCTGGCACCCACCCACACCTGCGGCGAGCTGGACCTGGGC 5268
QY 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
DB 5269 AGCAGCATCTGACAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCAATCCACC 5328
QY 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
DB 5329 TGCAGATACAGACCTTGATTTAGAAATGCACAGATTACCCCTGGTGTGACGACCTC 5388

```
Qy 738 TtpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 5389 TGGAGAAATCGTCATCGCCACCGTAACCTTAAGATCCCTCACTTGGGAGGACCCACCTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGGAAGAGAGATGTAAAGATGCGGTGTGAAGCCTTAAACACCCCAAAATGTTGTGGAG 5508
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATGTCTGTGGATGTGACCATGCTGTACCTGAAGATCTCCCAATC 5568
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTAGGACCTCCCGAGCTGAAATCTCTGAGCCTGGCAGGAAACAGGTGACAGACCG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCCTCTCAGTGATGCTGTGAGAGTCTCCAGTGGCCCTGCAGAAGCTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGGAGGACTGTGGCATCACAGCCACGGGTTCGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACTGTGCTTATCCAAACACACCTGGGGACAGAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTCGATCATGAGGCTTCCCACTGTAGTCTGTCAGAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTTGACACGGCTGGCTGTGTTCTTCTGCACTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGCACTTGAGCTTAGCATGAACTTGTGGAGCAATGGCGTGAACCTTCTGTGCG 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuLeuValLysCysHisLeu 957
Db 5989 GAGGTTCATGAGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTATCTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValLysSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCGCTGCTGTGAGAGTCTGTCTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuLysCysGluLeu 997
Db 6109 CTGGATCTCAGGCAATGCTTGGGTGACGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAACAGTGTCTTCAGCAGAGCTCGGGTTGAAGGAGCATGTGGACTGACTTCTGAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuLysSerAlaPheAlaCysPro 1057
Db 6289 GTGCAGATAATCTTCAGTCCCAAGAAATGATGAGCTGTGTGCGCCCTTTGCGCTGTGCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTAATTACATTAATTGGC 6372
```

RESULT 8

```
AY054986 3885 bp mRNA linear PRI 02-APR-2002
LOCUS Homo sapiens maternal-antigen-that-embryos-require protein (MATER)
```

```
mRNA, complete cds.
ACCESSION AY054986
VERSION AY054986.1
KEYWORDS GI:19882272
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3885)
AUTHORS Tong,Z.B., Bondy,C.A., Zhou,J. and Nelson,L.M.
TITLE A human homologue of mouse Mater, a maternal effect gene essential
for early embryonic development
JOURNAL Hum. Reprod. 17 (4), 903-911 (2002)
MEDLINE 21922687
PUBMED 11925379
REFERENCE 2 (bases 1 to 3885)
AUTHORS Tong,Z.-B., Bondy,C.A. and Nelson,L.M.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD,
NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA
```

FEATURES

```
Location/Qualifiers
source 1..3885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
gene 1..3885
/genes="MATER"
CDS 1..3603
/genes="MATER"
/notes="oocyte-specific protein; similar to Mus musculus
Mater encoded by GenBank Accession Numbers AF074018 and
AF143559-AF143573"
/codon_start=1
/product="maternal-antigen-that-embryos-require protein"
/protein_id="AAL15549.1"
/db_xref="GI:19882273"
/tranlations="MKVAGGLEIGAAALISASPRALVTLSTGPTCSILPKNPLPPONL
SSQPCIMEGDKSLTSSYGLQWCLVELDEKPEFTPKELLKKSSSTSCSIQPEIE
NANVECLALLHBYIGASLAWATISIFENNNRUTLSEKARDKDDKRRSPDPAITMD
QGSKEKPGISOAQVQDSATAETKEQEIISOAMEQEGATAETEQEISOAMEQEG
TAETEQEGGTDYKSHVMTKFAEEDVRRSFENTADWFMQTLGAFQSDRWG
FRRTVVLHGKSGIKSALARRIVLQWAGGLYQGMFSYVFFLPVREMQRKSSVTE
FISREMPDSQAPVTEIMSRPERLLFIIDGDDIGSVLNNDTKLCKWAEKQPPETLIR
SLRKVLLPESFLIVTRDVGTEKIKSEVSVPRYLLVIRGSGEQRIHLLEIRGIEHQ
QTGLKRALNNRELLDQCQVAVSLICVALQIQDVVSVAVFNFQTLTGHAAPFH
QLTPRGVVRNCLNEERVLKRFRCMAVGVNWRKSVFDDGLMVGQGLSGESLRALFH
MNTLLPDSHCEYYTTFHLSQDFCAALYYVLEGLETEPALCPLYVEKTRKSWELKQ
GFHISLMKRFLGLVSDVRRLPVLGCPVLGVKQKLLHWVSLGQGPKNATTPG
DTLDAPHCLPETODKEFVRLALNSFORVWLPINQNLDLIASFCLQHCPCYLRKTRDV
KGIFFPDESACPVVPLWNRDXTLIEEQWEDFCMLGTHPHLRQLDLGSSILTERAM
KTICAKLRHPTCKIQTLMFNAQITPGVQHLWRIVMANRNLSRLNGGTHLKEDVRM
ACBALHPKCLLESRLDCCGLTHACYLKIQLTTSPLSKLSLAGNKVTDQGVTP
SDALRYSQCALQKLILEDCCGTTATGQSLASVNSRSLTHCLSNNSLGNVGLLC
RSMRLPHCSLQRLMLNQCLDRTAGCGSLALAMGNSWLTHLSLMSNPVEDONGVKLCE
VMREPSCHLDLLELVKCHLTAAACESLSCVISRSRLKSLDLTDNALGDGVAAKCEG
LAKNSVFLRLGLKACGLTSDCEALSLASCNRHLTSLNLVQNPSFKGMKLCASF
ACPTSNLIQILGLMKWQIPVQIRKLESEVQLKPRVVIDGSWHSDDEDDRTWKN"
```

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3885
Score:	616.00	Matches:	1002
Percent Similarity:	99.21%	Conservative:	0
Best Local Similarity:	99.21%	Mismatches:	4
Query Match:	45.83%	Indels:	8
DB:	9	Gaps:	0

US-10-066-521-6 (1-1344) x AY054986 (1-3885)

```
Qy 98 IleserGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCACAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAGAACAAGAAATT 627
```

QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
Db 628 TCACAGCTATGGAAACAGAGAGGTGCCACACAGCAGAGACAGAGAACATGGGA 687
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTACAAGATCAGTGATGACCAAAATTCGCTGAGGAGGAGATGTA 747
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTCGTAGTTTTGAAAAACATCGCTGACTGGCGGGAATGCAAAAGTTCGCTGGCTGCT 807
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 808 TTTGANTTCAGACCGGTGGGGCTTCGGCCCTCGCACGGTGTTCGCAACGGAAATGCAGGA 867
QY 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db 868 ATTGGGAATCGGCTTAGCCAGAGGATCGTGCTGTGCTGGCGCAAGGTGGACTCTAC 927
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlyGlu 237
Db 928 CAGGGAATGTTCTCTACGCTCTTCTCTCCCGCTTAGAGAGATGCAGCGGAAGAGAG 987
QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrClu 257
Db 988 AGCAGTGTCAAGAGATTTCATCTCAGGAGGTGGCCAGACTCCCAAGGCTCCGGTGACCGAG 1047
QY 258 IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspLeuGlyLys 277
Db 1048 ATCATGTCGCCACCAAGAGCTGTGTTTCATATTGACGGTTTCGATGACCTGGGCTCT 1107
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 1108 GTCTCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGCCCTCCGTTCA 1167
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACGACGTCTGCTGAGGAAGTCTGCTGCTCCCTGAGTCCCTTCCTGATCGTCACCGTC 1227
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGCGTGGGCAACAGAAAGCTCAAGTCAGAGGTGCTGTCTCCCGTTACCTGTTAGTT 1287
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGGAACAAAGAATCCACTTGCTCTTGAGCCGGGATTTGGTGAGCAT 1347
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysGly 377
Db 1348 CAGAAGACACAAGGGTTCGCTGCGCATCAT-CAACAACCGTGAGCTGCTCAGCAGTCCCA 1406
QY 377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuAspValValGlyGly 397
Db 1407 GGTGCCCGCGTGGGCTCTCATCTCGCTGGCCCTCGACGCTGCAGACGTGGTGGGGGA 1466
QY 397 userValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db 1467 GAGCGTCGCCCCCTTCAACCAACAGCTCACAGGCTGCAGCGCTGTCAGCCGCTTTTGC-GTTTCATC 1525
QY 417 InLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db 1526 AGCTACACCCCTCGAGGGGTGGTCCGGCGCTGCTCAATCTGGAGGAAGAGTTGTCTCTGA 1585
QY 437 ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db 1586 AGCGCTTCGCGGTATGGCTGTGAGGGAGTGTGGAAATAGGAAGTCAGTGTTGATGGTG 1645
QY 457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db 1646 ACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGAACA 1705

QY 477 leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA 497
Db 1706 TCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCAGTCTCAGG 1765
QY 497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db 1766 ACTTCTGTGCGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCC 1825
QY 517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db 1826 CTCTGTACGTGTGAGAAGACAAAGAGGTCCATGAGCTTAAACAGGAGGCTTCCATATCC 1885
QY 537 isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProL 557
Db 1886 ACTCGCTTTGGATGAAGCGTTTCTTGTGTGGCTCTGTGAGCGAAGACGTAAAGAGGCCAC 1945
QY 557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV 577
Db 1946 TGGAGGTCTGTCTGGGCTGTCCCGTTCCTGGGGGTGAAGCAGAGACTTCTGCACCTGGG 2005
QY 577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db 2006 TCTCTCTGTGGGTACGACAGCTTAATGCCACCAACCCAGGAGACACCTGGAGCGCTTCC 2065
QY 597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db 2066 ACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACAGACTTCCAAG 2125
QY 617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db 2126 AAGTGTGGCTTCGATTACACAGAACCTGACTTGATAGCATCTTCTTCTGCTCCAGC 2185
QY 637 isCysProTyrLeuArgLysIleArgValaspValLysGlyIlePheProArgAspGluS 657
Db 2186 ACTGTCCGTATTTGCGGAAAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGT 2245
QY 657 erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db 2246 CCCTCAGGAGCATGCTCTGTGCTCCCTCTATGGATCGGGATAGACCCCTCATTTGAGGAGC 2305
QY 677 InTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db 2306 AGTGGGAAGATTTCTGCTCCATGCTTGGCCACCCACCCACCTGCGCGCAGCTGGACCTGG 2365
QY 697 lysSerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db 2366 GCAGCAGCATCTCTGACAGAGCGGCGCATGAAGACCCCTGTGTGCCAAGCTGAGGATCCCA 2425
QY 717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db 2426 CCTGCAAGATACAGACCCCTGATGTTTAGAAATGCAGATTTACCCCTGGTGTGCCAACACC 2485
QY 737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db 2486 TCTGGAGAATCGTGTATGCGCAACCGTAACCTAAGATCCCTCAACTTGGGAGGCCACCCACC 2545
QY 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAGAGGATGTAAGGATGGCGTGTGAAGCCCTTAAACACACCCCAAAATTTGTTGG 2605
QY 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLeuValIleSerGlnI 797
Db 2606 AGTCTTTGAGGCTGGGATGCTGTGGATTCACCATCCCTGTTTACCTGGAAGATCTCCAAA 2665
QY 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCTTACGACCTTCCCCAGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACC 2725
QY 817 InGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGGAGTAAC-GCCTCTCAGTGATGCTTTGAGGGTCTCCAGTGCCTCCCTGCAAGAGCTG 2784
QY 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856

QY 258 IleMetSerArgProGluAtrqLeuLeuPheIleleaspGlyPheAspAspLeuGlySer 277
Db 1048 ATCATGTCCCGACACAGAAAGGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 1107
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db 1108 GTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGCCTCCGTTCAAC 1167
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1168 CTCATACGACGTCTGCTGAGAAAGGTCCTGCTCCCTGAGTCTCTCTGATCGTCAACGCTC 1227
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db 1228 AGAGACGTGGGCACAGAAAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTAGTT 1287
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 1288 AGAGGAATCTCCGGGGAACAAAGAAATCCACTTGTCTCTTGAGCGCGGATTTGGTGACAT 1347
QY 358 GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuLeuAspGlnCysGl 377
Db 1348 CAGAAGACACAAAGGTTGCGTSCGATCAT-CAACAACCGTGAGTGTCTCGACACAGTCCCA 1406
QY 377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGl 397
Db 1407 GGTGCCCGCCGTGGGCTCTCATCTGGTGGCCCTGCAGCTGCAGACGTGGTGGGGA 1466
QY 397 uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG 417
Db 1467 GAGCGTCGCCCCCTTCAACCAACCGCTCACAGGCTGCACGCCCTTTTGC-GTTTCATC 1525
QY 417 lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db 1526 AGCTCACCCCTCGAGGGCTGTGTCGGGCGCTCTCAATCTGGAGGAAGAGTTGTCTCTGA 1585
QY 437 ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db 1586 AGCGCTTCTGCGGTATGCTGTGAGGAGGTGGGAATAGGAAGTCAAGTGTGTATGGTG 1645
QY 457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db 1646 ACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTCCTCTGTTTCACATGAACA 1705
QY 477 leLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnA 497
Db 1706 TCCTTCTCCACAGACGCACTGTGAGGAGTACTACACTTCTTCCACTCAGTCTCCAGG 1765
QY 497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db 1766 ACTTCTGTGCCGCTTGTACTACTGTGTAGAGGCTGGGAATCGAGCCAGCTCTCTGCC 1825
QY 517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db 1826 CTCGTGTACGTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGCGAGGCTTCCATATCC 1885
QY 537 isSerLeuTrpMetLysArgPheIleuPheGlyLeuValSerGluAspValArgArgProL 557
Db 1886 ACTCGCTTTGGATGAAGCGTTTCTGTTGGCCCTCGTGAGCGAAGACGTAAAGGAGGCCAC 1945
QY 557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV 577
Db 1946 TGGAGGTCCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAACGAGAAGCTCTGCACTGGG 2005
QY 577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db 2006 TCTCTCTGTTGGGTGAGCAGCCTTAATGCCACACCCCGAGGAGACACCTCGAGCGCTTCC 2065
QY 597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db 2066 ACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAACACAGCTTCCAAG 2125

QY 617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db 2126 AAGTGTGGCTTCGATTTAACCAAGACCTGAGCTTGTATAGATCTTCTTCTGCTCCAGC 2185
QY 637 isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS 657
Db 2186 ACTGTCCGTATTTCGGGAAAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATGAGT 2245
QY 657 erAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db 2246 CCCTGAGGCGATCTCTGTGTCCCTCTATGGATGGGGATAAGACCTCTCATTTGAGGAGC 2305
QY 677 lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db 2306 AGTGGGAGAGATTTCGTCTCATCTTTGGCACCCACCCACCTGCGGCGAGCTGGACTGG 2365
QY 697 lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db 2366 GCAGCAGCATCTCTGACAGAGCGGCCCATGAAGACCTGTGTGTGCCAAGCTGAGGCATCCCA 2425
QY 717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db 2426 CCTGCAAGATACAGACCTCGATGTTTTAGAAATGCACAGATTACCCCTGGTGTGCCACACC 2485
QY 737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db 2486 TCTGGAGATCTGTATGGCCACCGTAACCTAAGATCCCTCAACTTTGGGAGGCACCCACC 2545
QY 757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db 2546 TGAAGGAAGAGGATGTAAGGATGGCTGTGAAGCTTAAAAACACCCAAAAATGTTGTGG 2605
QY 777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db 2606 AGTCTTTGAGGCTGGATGCTGTGGAATGACCATGCCCTGTTACCTGAAGATCTCCAAA 2665
QY 797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db 2666 TCCTTAGACCTCCCGACGCTGAATCTCTGAGCTTGGCAGGAAACAAGGTGACAGACC 2725
QY 817 lnGlyValMet-ProfenSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu 836
Db 2726 AGGGAGTAAAC-GCCTCTCAGTGATGCTTGAAGGCTCTCCAGTGCCTGCAAGAAAGCTG 2784
QY 837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db 2785 ATACTGGAGACTGTGGCATCACAGCCAGGGTTGCCAGAGTCTGGCCTCAGCCCTCGTC 2844
QY 857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db 2845 AGCAACGGGAGCTTGACACACCTTGCTTATCCAAACAACAGCTGGGGAACGAAGGTGTA 2904
QY 877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db 2905 AATCTACTGTGTGCATCCATGAGCTTCCCACTAGTCTGACAGAGCTGATGCTGAAT 2964
QY 897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db 2965 CAGTGGCACCTGGACAGCGCTGGCTGTGG-TTCTCTTGACACTTGGCTTATGGGTAACTC 3023
QY 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe 936
Db 3024 ATGGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAGACAATGGCGTGAAGCTTCT 3083
QY 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi 956
Db 3084 GTGCGAGGTTCATGAGAGAACCATCTTGTCTATCTCTCCAGGACCTGGAGTTGGTAAAGTCA 3143
QY 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
Db 3144 TCTACCGCCGCGCTGTGTGAGAGTCTGCTCTGTGTATCTCGAGGAGCAGACACCTGAA 3203
QY 976 sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValValAlaAlaLeuCysGluGl 996

Db	3204	GAGCCTGGATCTCACGGACAATCCCTGGGTGACGGTGGGGTTGCTGCGCTGGCAGGG	3263		Query Match: 39.66%
Qy	996	YLeuLysGlnLysAasnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe	1016		Indels: 0
Db	3264	ACTGAACGAAAGAACAGTGTCTGACGACACTCGGGTTGAAGCAATGGACTGACTTC	3323		Gaps: 0
Qy	1016	rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAasnArgHisLeuThrSerLeuAs	1036		US-10-066-521-6 (1-1344) x AC011470 (1-157141)
Db	3324	TGATTGCTGTGAGGCACTCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTAAA	3383		Qy 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu
Qy	1036	nLeuValGlnAasnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCys	1056		Db 42000 GGACATGGAGGTGACACATGGGACTTACAAAGAGTCACGTGATGACCAATTCGCTGAGGAG
Db	3384	CCTGGTGCAAAATACTTCAGTCCCAAGAAATGATGAAGCTGTGTTCGGCCCTTTGCCCTG	3443		Qy 155 GluAspValArgArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeu
Qy	1056	sProThrSerAasnLeuGlnIlelleGlyLeuTrpLysTrpGlnTyrProValGlnIleAr	1076		Db 41940 GAGGATGTACGTGTGTAGTTTTTGAATAACACTCTCTGCTACTGGCCGAAATGCAACCGTTG
Db	3444	TCCACGCTCTAATTACAGATAATTTGGCTGTGGAATGGCAGTACCTCTGCAATAAAG	3503		Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgTrpValValLeuHisGly
Qy	1076	gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi	1096		Db 41880 GCTGGTCTTTGATTTCAGACCGGTGGGGCTTCGGCCCTCGCACGGTGGTTCTTCGACCGA
Db	3504	GAAGCTGCTGGAGGAAGTGCAGTACTCAAGCCCCGAGTCGTAATTGACGGTAGTTGGCA	3563		Qy 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly
Qy	1096	sSerPheAspGluAspAspArg 1103			Db 41820 AAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTGTCTGTGGCGCAGGT
Db	3564	TTCTTTTGTGAAGATGACCGG 3585			Qy 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg
RESULT 10					Db 41760 GGACTCTATACAGGGAATGTTCTCTCTACGTCTTCTTCTCTCCCGTTAGAGATGCAGCGG
AC011470/c					Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro
LOCUS	AC011470	157141 bp	DNA	linear	Db 41700 AAGAAGGAGACGAGTGTACAGAGTTCTATCTCCAGGGNGTGGCCAGACTCCAGGCTCCG
DEFINITION	Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.				Qy 255 ValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAsp
ACCESSION	AC011470				Db 41640 GTGACGGAGATCATGTCCCGACACAGAAAGGCTGTGTTCATCATTCATGACGGTTTCGATGAC
VERSION	AC011470.5	GI:9211204			Qy 275 LeuGlySerValLeuAasnAasnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro
KEYWORDS	HTG.				Db 41580 CTGGGCTCTGTCTTCAACATGACACAAAGCTCTGCAAAAGACTGGCGTGAAGACGAGCT
SOURCE	Homo sapiens (human)				Qy 295 ProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluLysPheLeuIle
ORGANISM	Homo sapiens				Db 41520 CCGTTCACCTCTATACGCACTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCTCTCTGATC
REFERENCE	1. (bases 1 to 157141)				Qy 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.				Db 41460 GTCACCGTCAGACGCTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCTCCCGTTAC
TITLE	Direct Submission				Qy 335 LeuLeuValArgGlyIleSerGlyGlnArgIleHisLeuLeuLeuGluArgGlyIle
JOURNAL	Unpublished				Db 41400 CTGTTAGTTAGAGAATCTCCGGGGAACAAAGAATCCACTTGCTCTTGTGACGCGGGATT
REFERENCE	2. (bases 1 to 157141)				Qy 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAenAenArgGluLeuLeuAsp
AUTHORS	DOE Joint Genome Institute.				Db 41340 GGTGAGCATCAGAGACACAAAGGGTTGGTGCGATCATGAACAACCGTGAGCTCTCGAC
TITLE	Direct Submission				Qy 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal
JOURNAL	Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				Db 41280 CAGTGCCAGGTGCCCGCGTGGGTCTCTCATCTGCTGGCCCTGCAGCTGCAGGCGTG
REFERENCE	3. (bases 1 to 157141)				Qy 395 ValGlyGluSerValAlaPropheAasnGlnThrLeuThrGlyLeuHisAlaAlaPheVal
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.				Db 41220 GTGGGGGAGAGCGTCGCCCTTCAACCAACCGCTCAAGGCTTCGACCGCGCTTTTGTG
TITLE	Direct Submission				Qy 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAenLeuGluGluArgVal
JOURNAL	Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				Db 41160 TTTTCATCAGCTCACCCCTCGAGGGCGTGTCCGGCGCTGTCTCTCAATCTGGAGAAAGAGTT
COMMENT	On Jul 15, 2000 this sequence version replaced gi:7690109. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.sngc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.1. Location/Qualifiers 1. .157141 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /clone="CTC-490M10"				Qy 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAenArgLysSerValPhe
FEATURES	source				Db 41100 GTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGTGGATAGGAAGAGTT
ORIGIN					Qy 455 AspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis
Alignment Scores:					Db 41040 GACGTCGACGACCTCATGTTCAGAGGACTCGGGGAGTCTGAGTCTCGGTCTCTGTTTCAC
Pred. No.:	0	Length:	157141		
Score:	533.00	Matches:	533		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		


```
QY 475 MetAenlleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSer 494
Db 40980 ATGAACATCTCTTCCAGACAGCCACTGTGAGGAGTACTACCTTCTCCACCTCAGT 40921
QY 495 LeuGlnAspPheCysAlaLeuTyrThrValLeuGluGlyLeuGluLeuLeuProAla 514
Db 40920 CTCACGAGACTTCTGTGCGCCCTTGTTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCT 40861
QY 515 LeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
Db 40860 CTCCTGCCCTCTGTAGCTGTGAGAAGCAAAAGAGGTCATGGAGCTTAAACAGGACGGCTTC 40801
QY 535 HisIleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
Db 40800 CATATCCACTGCTTTGGATGAGCGTTTCTGTGTGGCCTCGTGAGGAGAGCTAAGG 40741
QY 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
Db 40740 AGGCCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCTCCCTGGGGGTGAAGCAGAGACTTCTG 40681
QY 575 HisTrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAsp 594
Db 40680 CACTGGGTCTCTCTGTGGGTGAGCAGCCTAATGCCACACCCAGGAGACACCCCTGGAC 40621
QY 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
Db 40620 GCCTTCCACTGTCTTTTCGAGACTCAGCAACAAGAGTTGTTCGCTGGCATTAACAGC 40561
QY 615 PheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCys 634
Db 40560 TTCCAAGAAGTGTGGCTCCGATTAATACCAGAACCTGGACTTGATAGCATCTTCTCTGCG 40501
QY 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
Db 40500 CTCAGACACTGCTCGCTATTTTCGGGAAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGA 40441
QY 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667
Db 40440 GATGAGTCCGTGAGGATGCTCTGTGGTCTCTATGG 40402

RESULT 11
AC024580 193609 bp DNA linear PRI 21-DEC-2001
LOCUS Homo sapiens chromosome 19 clone CTD-2621117, complete sequence.
DEFINITION AC024580
ACCESSION AC024580
VERSION AC024580.6 GI:17975240
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Unpublished
REFERENCE 2 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
TITLE Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Submitted (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 193609)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Dec 21, 2001 this sequence version replaced gi:16905144.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
```

```
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.9.
FEATURES
    source
        1. 193609
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /clone="CTD-2621117"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 193609
Score: 533.00 Matches: 533
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.66% Indels: 0
DB: 9 Gaps: 0
US-10-066-521-6 (1-1344) x AC024580 (1-193609)
QY 135 GlyHisGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 154
Db 22258 GGACATGGAGGTGACATCGGACTACAGAGTCACGTGATGACCAATTCGCTGAGGAG 22317
QY 155 GluAspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeu 174
Db 22318 GAGGATGTACGTGCTAGTATTTTGAATAACACTGCTGCTGACTGGCCGGAATGCAACGTTG 22377
QY 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
Db 22378 GCTGGTGCTTTTGATTCAGACCGGTGGGCTTCGCGCTCGCACGGTGGTTCGTCACGA 22437
QY 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGly 214
Db 22438 AAGTCAGGAATTTGGGAATTCGGCTCTAGCCAGAGAGATCGTCTGTGCTGGGCGCAAGT 22497
QY 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
Db 22498 GGACTCTACAGGGAATGTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGCAGCGG 22557
QY 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
Db 22558 AAGAAGGAGCAGCTGTCACAGAGTTCTCTCAGGGAGTGGCCAGACTCCCAGGCTCCG 22617
QY 255 ValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAsp 274
Db 22618 GTGACCGAGATCATGTCCGACACAGAAAGGCTGTGTTTCATCATTTGACGTTTCGATGAC 22677
QY 275 LeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnPro 294
Db 22678 CTGGGCTCTGTCTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCT 22737
QY 295 ProPheThrIleLeuArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIle 314
Db 22738 CCGTTCCACTCATACGAGTCTGTGAGAAAGGTCCTGCTCCCTGAGTCCCTTCTCTGATC 22797
QY 315 ValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyr 334
Db 22798 GTCAACGTCAGAGACGTGGGCACAGAAAGCTCAAGTCAGAGGTCTGTCTCCCCGTTAC 22857
QY 335 LeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIle 354
Db 22858 CTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCCACTTGTCTCTTTCGAGCGCGGATT 22917
QY 355 GlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAsp 374
Db 22918 GGTGAGCATCAGAAGACACAAGGTTTCGCTGGATCATCAACCAACCGTGCAGTTCGTCGAC 22977
QY 375 GlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspVal 394
```


Db 22978 CAGTCCAGGTCGCCGCGTGGGCTCTCTCATCTCGGTGGCCCTGCAGCTCGAGAGCTG 23037
Qy 395 ValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal 414
Db 23038 GTGGGGGAGAGGTGCGCCCTTCAACCAAGCTCACAGGCTGCAGCGCTTTGTG 23097
Qy 415 PheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgVal 434
Db 23098 TTTCATCAGCTCACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTCGAGAAAGATT 23157
Qy 435 ValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPhe 454
Db 23158 GTCCCTGAAGCGCTTCTGCGGTATGCTGTGAGGAGGTGTGGAATAGGAATGCTGTT 23217
Qy 455 AspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHis 474
Db 23218 GACGGTGACGACCTCATGTTCAAGGACTCGGGGAGTCTGAGCTCCGTCCTGTTTCAC 23277
Qy 475 MetAsnLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSer 494
Db 23278 ATGAACATCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTCCACCTCAGT 23337
Qy 495 LeuGlnAspPheCysAlaAlaLeuTyrThrValLeuGluGlyLeuGluLeuGluProAla 514
Db 23338 CTCAGGACTTCTGTGCGCGCTTGTACTACGTGTGTAGAGGCGCTGGAATCGAGCCAGCT 23397
Qy 515 LeuCysProLeuTyrValGlnLysThrLysArgSerMetGluLeuLysGlnAlaGlyPhe 534
Db 23398 CTCTGCCCTCTGTAGTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGCGAGGCTTC 23457
Qy 535 HisLeHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArg 554
Db 23458 CATATCCACTCGCTTTGGATGAAGCGTTCTTGTGTGGCTCGTGAGCGAAGCGTAAGG 23517
Qy 555 ArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeu 574
Db 23518 AGCCACCTGGAGGCTCTCTGCGGCTGTCCCGTTCCTCGGGGGTGAAGCAGAGCTTCTG 23577
Qy 575 HisTrpValSerLeuLeuGlyGlnProAsnAlaThrProGlyAspThrLeuAsp 594
Db 23578 CACTGGGTCTCTGTGTGGTTCAGCAGCTTAATGCCACCACCCAGGAGACACCTTGAC 23637
Qy 595 AlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSer 614
Db 23638 GCCTTCCACTCTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGTTGGCATTAAACAGC 23697
Qy 615 PheGlnGluValTrpLeuProLeuAsnGlnAsnLeuAspLeuLeuAlaSerSerPheCys 634
Db 23698 TTCCAAGAGGTGGCTTCCGATTAAACAGAACCTGGACTTGTATGATCATCTTCTCTGTC 23757
Qy 635 LeuGlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArg 654
Db 23758 CTCAGCACTGCTCGTATTTCCGGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGA 23817
Qy 655 AspGluSerAlaGluAlaCysProValValProLeuTrp 667
Db 23818 GATGAGTCCGCTGAGGATGCTCTGTGCTCCCTCTATGG 23856

RESULT 12
AX427588 LOCUS 1157 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from Patent WO0232955.
ACCESSION AX427588
VERSION AX427588.1 GI:21537711
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nelson, L.M. and Tong, Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 1 25-APR-2002;

GOVERNMENT OF THE UNITED STATES (US)

FEATURES

Location/Qualifiers

1. 1157
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 1157
Score: 385.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.65% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX427588 (1-1157)

Qy 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuPheIleLeuAspGly 271

Db 1 CAAGCTCCGGTGACGGAGATCATGTCCGACCAAGAGGCTGTGTTTCATCATTTGACGGT 60

Qy 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGlu 291

Db 61 TTGATGACCTGGGCTCTGTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG 120

Qy 292 LysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311

Db 121 AAGCAGCTCCGTTACCTCATACGAGTCTGTGAGGAAGGTCTGCTCCCTCCCTGAGTCC 180

Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331

Db 181 TTCTGATGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCT 240

Qy 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGlu 351

Db 241 CCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTGCCTCTTGAG 300

Qy 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371

Db 301 CGCGGAGATTGGTGGAGCATCAAGAAGCAACAAGGTTGGTCCGATCATGAACACCGTGAG 360

Qy 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391

Db 361 CTGCTGCACAGTCCAGCGTCCCGCGTGGGCTCTCTCATCTCGGTGGGCCCTGCAGCTG 420

Qy 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411

Db 421 CAGGACGTGGTGGGAGAGAGCGTCGCCCTTCAACCAACCGCTCACAGGCTGCACGCC 480

Qy 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431

Db 481 GCCTTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAG 540

Qy 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451

Db 541 GAAAGAGTTGCTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGGTGTGGAATAGGAAG 600

Qy 452 SerValPheAspGlyAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471

Db 601 TCAGTGTTCGACGGTGACGACCTCATGGTTCAGAGGACTCGGGGAGTCTGAGCTCCGTGCT 660

Qy 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrThrPhePhe 491

Db 661 CTGTTTCATCATGAACATCTCTTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTC 720

Qy 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIle 511

Db 721 CACTCAGTCTCCAGGACTTCTGTGCGCGCTTGTACTAGTGTAGAGGCGCTTGAATC 780

Qy 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531

Db 781 GAGCCAGCTCTCTGCTCTGTACGTTGAGAACAAAGAGGTCCTCATGGAGCTTAAACAG 840

QY 532 AlaGlyPheHisIleHisSerLeuTrpMetIysArgPheLeuPheGlyLeuValSerGlu 551
 Db 841 CGAGGCTTCCATATCCACTCGCTTGGATGAAGGTTCTTGTTGGCCCTCGTAGCGAA 900
 QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValIysGln 571
 Db 901 GACGTAGAGGCCACCTGGAGTCTCTGCTGGGCTGTCCGTTCCCTTGGGGGTGAACAG 960
 QY 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrProGlyAsp 591
 Db 961 AAGCTTCTGCACCTGGGCTCTCTCTGGGTGAGCAGCCTAATGCCACCCAGGAGAC 1020
 QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspIysGluPheValArgLeuAla 611
 Db 1021 ACCCTGGAGCGCTTCCACTGCTCTTTCGAGACTCAAGCAAGAAGTTGTTCGCTGCGCA 1080
 QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
 Db 1081 TTAACAGCTTCCAGAGAGTGGCTCCGATTACCCAGAACCTGGACTGTATAGCATCT 1140
 QY 632 SerPheCysLeuGln 636
 Db 1141 TCCTTCTGCTCCAG 1155

RESULT 13
 AC012107/c
 LOCUS Homo sapiens clone RP11-45K21, WORKING DRAFT SEQUENCE, 25 unordered
 DEFINITION
 AC012107
 AC012107.2 GI:7329252
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 167509)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens, clone RP11-45K21
 Unpublished
 2 (bases 1 to 167509)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
 Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 26, 2000 this sequence version replaced gi:6088020.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L1180
 Center clone name: 45_K_21
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads
 Chemistry: Dye-terminator Big Dye; 96% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 145437 bases at least Q40
 Consensus quality: 155496 bases at least Q30
 Consensus quality: 159832 bases at least Q20
 Insert size: 176000; agarose-fp
 Insert size: 165109; sum-of-contigs
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 4.1 in Q20 base.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 25 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1061: contig of 1061 bp in length
 * 1062 1161: gap of 100 bp
 * 1162 2381: contig of 1220 bp in length
 * 2382 2481: gap of 100 bp
 * 2482 3771: contig of 1290 bp in length
 * 3772 3871: gap of 100 bp
 * 3872 4949: contig of 1078 bp in length
 * 4950 5049: gap of 100 bp
 * 5050 6119: contig of 1070 bp in length
 * 6120 6219: gap of 100 bp
 * 6220 8023: contig of 1804 bp in length
 * 8024 8123: gap of 100 bp
 * 8124 10555: contig of 2432 bp in length
 * 10556 10655: gap of 100 bp
 * 10656 12768: contig of 2113 bp in length
 * 12769 12868: gap of 100 bp
 * 12869 15722: contig of 2854 bp in length
 * 15723 15822: gap of 100 bp
 * 15823 17959: contig of 2137 bp in length
 * 17960 18059: gap of 100 bp
 * 18060 21771: contig of 3712 bp in length
 * 21772 21871: gap of 100 bp
 * 21872 24539: contig of 2668 bp in length
 * 24540 24639: gap of 100 bp
 * 24640 28144: contig of 3505 bp in length
 * 28145 28244: gap of 100 bp
 * 28245 31503: contig of 3259 bp in length
 * 31504 31603: gap of 100 bp
 * 31604 35858: contig of 4255 bp in length
 * 35859 39610: contig of 100 bp
 * 39611 39710: gap of 100 bp
 * 39711 42759: contig of 3049 bp in length
 * 42760 42859: gap of 100 bp
 * 42860 48014: contig of 5155 bp in length
 * 48015 48114: gap of 100 bp
 * 48115 52263: contig of 4149 bp in length
 * 52264 52363: gap of 100 bp
 * 52364 56374: contig of 4011 bp in length
 * 56375 63102: contig of 6628 bp in length
 * 63103 63202: gap of 100 bp
 * 63203 72687: contig of 9485 bp in length
 * 72688 72787: gap of 100 bp
 * 72788 97799: contig of 25012 bp in length
 * 97800 97899: gap of 100 bp
 * 97900 134325: contig of 36426 bp in length
 * 134326 134426: gap of 100 bp
 * 134426 167509: contig of 33084 bp in length.

FEATURES
 source
 1. 167509
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-45K21"

/clone_lib="RPCI-11 Human Male BAC"

1..1061
/note="assembly_fragment"
1162..2381
/note="assembly_fragment"
2482..3771
/note="assembly_fragment"
3872..4949
/note="assembly_fragment"
5050..6119
/note="assembly_fragment"
6220..8023
/note="assembly_fragment"
8124..10555
/note="assembly_fragment"
10656..12768
/note="assembly_fragment"
12869..15722
/note="assembly_fragment"
15823..17959
/note="assembly_fragment"
18060..21771
/note="assembly_fragment"
21872..24539
/note="assembly_fragment"
24640..28144
/note="assembly_fragment"
28245..31503
/note="assembly_fragment"
31604..35858
/note="assembly_fragment"
35959..39610
/note="assembly_fragment"
39711..42759
/note="assembly_fragment"
42860..48014
/note="assembly_fragment"
48115..52263
/note="assembly_fragment"
52364..56374
/note="assembly_fragment"
56475..63102
/note="assembly_fragment"
63203..72687
/note="assembly_fragment"
clone_end:T7
vector_side:right
72788..97799
/note="assembly_fragment"
97900..134325
/note="assembly_fragment"
134426..167509
/note="assembly_fragment"
clone_end:SP6
vector_side:left

ORIGIN

Alignment Scores:
Pred. No.: 2,87e-315 Length: 167509
Score: 300.00 Matches: 300
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.32% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x AC012107 (1-167509)

Qy 368 AnAsnArgGluLeuLeuAspGlnCysGlnValProAlaValGlySerLeuLeuCysVal 387
Db 12664 AACACCGTGAGCTCTCTGACACAGTCCAGGTGCCCGCGGTGGGCTCTCTCATCTGGGTG 12605
Qy 388 AlaLeuGlnLeuGlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThr 407

Db 12604 GCCTCGAGCTGCAGGACGTGGTGGGGAGAGCGCTCGCCCCCTTCAACCAACGCTCACA 12545
Qy 408 GlyLeuHisAlaAlaPheValPheHisGlnLeuThrProArgGlyValValArgCys 427
Db 12544 GGCCTGCACGCCGCTTTTGTTCATCAGCTCACCCCTCGAGGCGGTGTCGGCGGTGT 12485
Qy 428 LeuAsnLeuGluArgValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 447
Db 12484 CTCAACTCGGAGGAAAGAGTTGTCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTG 12425
Qy 448 TrpAsnArgGlySerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSer 467
Db 12424 TGGAAATAGGAAGTCAGTGTTCATGAGAGACCTCATGGTTCAAGGACATCGGGGAGTCT 12365
Qy 468 GluLeuArgAlaLeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyr 487
Db 12364 GAGCTCGTGTCTGTGTTCATGATGAACATCTTCTCCAGACGCCACTGTGAGAGTAC 12305
Qy 488 TyrThrPhePheHisLeuSerLeuGlnAspPheCysAlaLeuLeuTyrTyrValLeuGlu 507
Db 12304 TACACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCCTTGTACTACGTGTAGAG 12245
Qy 508 GlyLeuGluLeuGluProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 527
Db 12244 GGCCTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACCTTGAGAGACAAAGAGTCCATG 12185
Qy 528 GluLeuLeuGlnAlaGlyPheHisIleHisSerLeuTyrMetLeuLeuLeuLeuLeuLeu 547
Db 12184 GAGCTTAAACAGGAGGCTTCCATATCCATCTCGCTTGGATGAAGCGCTTCTGTGGTGGC 12125
Qy 548 LeuValSerGluAspValArgProLeuGluValLeuLeuLeuGlyCysProValProLeu 567
Db 12124 CTCGTGAGCGAAGACGTAAAGAGGCCACTGGAGGCTCTGTGGGCTGTCCGCTTCCCTG 12065
Qy 568 GlyVallySGlnLeuLeuLeuHisTyrValSerLeuLeuGlyGlnGlnProAsnAlaThr 587
Db 12064 GGGGTGAAGCAGAGCTTCTGCACCTGGGCTCTCTGTGGGTTCAGCAGCCCTAATGCCACC 12005
Qy 588 ThrProGlyAspThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPhe 607
Db 12004 ACCCAGAGACACCTCGAGCGCTTCCACTGTCTTTTCAGAGCTCAAGACAAAGAGTTT 11945
Qy 608 ValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAsp 627
Db 11944 GTTCGCTTGGCATTAAACAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAACCTGAC 11885
Qy 628 LeuIleAlaSerSerPheCysLeuGlnHisCysProTyrLeuArgLysIleArgValAsp 647
Db 11884 TTGATAGCATCTTCTTCTCCCTCCAGCACTGCCGTATTTGCGGAAAATTCGGGTGGAT 11825
Qy 648 VallySGlyIlePheProArgAspGluSerAlaGluAlaCysProValValProLeuTyr 667
Db 11824 GTCAAGGGATCTTCCAGAGATGAGTCGCTGAGGAGTGTCTGTGGTGGCTCTATGG 11765

RESULT 14
AX427590
LOCUS AX427590 1075 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from Patent WO0232955.
ACCESSION AX427590
VERSION AX427590.1 GI:21537712
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 3 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
1..1075
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.87e-203 Length: 1075
Score: 196.00 Matches: 252
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AX427590 (1-1075)

```
QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
   |||||
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGGAGCTGTGACACACCTGTGCCTATCCAAACAGC 61
   |||||
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
   |||||
Db 62 CTGGGGAACGAGGTGTAAATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTG 121
   |||||
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
   |||||
Db 122 CAGAGGCTGATGCTGAATCAGTCCACCTGACACGCGCTGGCTGTGG-TTCTCTTGCACT 180
   |||||
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
   |||||
Db 181 TGCCTGTATGGGTAACTCATGGCTGACGCACTGAGCCTTGTAGCATGAACCTGTGGAGA 240
   |||||
QY 930 pAsnGlyVallyLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
   |||||
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACCACTTGTCTATCTCCAGGACCT 300
   |||||
QY 950 uGluLeuVallyCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
   |||||
Db 301 GGAGTTGGTAAATGTCATCTCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTC 360
   |||||
QY 970 rArgSerArgHisLeuLeuSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValVa 990
   |||||
Db 361 GAGAGCAGACACCTGAGAGCCCTGGATCTCACGGACAATGCCCTGGGTGACGCTGGGGT 420
   |||||
QY 990 lAlaAlaLeuCysGluGlyLeuLeuGlnLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
   |||||
Db 421 TGCTGGCTGTGCGAGGACTGAGACCAAGAACAGTGTCTGACGAGACTCGGGTTGAA 480
   |||||
QY 1010 sAlaCysGlyLeuThrSerAspCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
   |||||
Db 481 GGCATGTGACTGACTTCTGATTGCTGTGAGGCACCTCTCCTTGGCCCTTTCTCTGCAACCG 540
   |||||
QY 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAspPheSerProlysglyMetMetLysLe 1050
   |||||
Db 541 GCATCTGACCACTAAACCTGGTGCAGATAAATCTCAGTCCCAAGGAATGATGAAGCT 600
   |||||
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070
   |||||
Db 601 GTGTTGGCCCTTGGCTGTCACAGTCTAACTTACAGATATTGGGCTGTGGAATGGCA 660
   |||||
QY 1070 nTyrrProValGlnIleArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValVa 1090
   |||||
Db 661 GTACCTGTGCAATAGGAAGCTGTGGAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGT 720
   |||||
QY 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
   |||||
Db 721 AATTGACGGTAGTTGGCATTCTTTTGTATGAGATGACCGG 760
   |||||
```

RESULT 15

AC012107 LOCUS 167509 bp DNA linear HTG 28-MAR-2000
DEFINITION Homo sapiens clone RP11-45K21, WORKING DRAFT SEQUENCE, 25 unordered
pieces.
AC012107 AC012107
VERSION AC012107.2 GI:7329252

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 167509)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens, clone RP11-45K21

Unpublished

2 (bases 1 to 167509)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Gage, M.,
Galan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Holland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, K., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 26, 2000 this sequence version replaced gi:6088020.
All repeats were identified using RepeatMasker:

Smit, A. P. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1180

Center clone name: 45_K_21

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-primer-amersham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 145437 bases at least Q40

Consensus quality: 155496 bases at least Q30

Consensus quality: 159832 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 165109; sum-of-contigs

Quality coverage: 3.8 in Q20 bases; agarose-fp

Quality coverage: 4.1 in Q20 base.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1061: contig of 1061 bp in length
* 1062 1161: gap of 100 bp
* 1162 2381: contig of 1220 bp in length
* 2382 2481: gap of 100 bp
* 2482 3771: contig of 1290 bp in length
* 3772 3871: gap of 100 bp
* 3872 4949: contig of 1078 bp in length
* 4950 5049: gap of 100 bp
* 5050 6119: contig of 1070 bp in length
* 6120 8023: contig of 1804 bp in length
* 8024 8123: gap of 100 bp
* 8124 10555: contig of 2432 bp in length
* 10556 10655: gap of 100 bp

* 10656 12768: contig of 2113 bp in length
* 12769 12868: gap of 100 bp
* 12869 15722: contig of 2854 bp in length
* 15723 15822: gap of 100 bp
* 15823 17959: contig of 2137 bp in length
* 17960 18059: gap of 100 bp
* 18060 21771: contig of 3712 bp in length
* 21772 21871: gap of 100 bp
* 21872 24539: contig of 2668 bp in length
* 24540 24639: gap of 100 bp
* 24640 28144: contig of 3505 bp in length
* 28145 28244: gap of 100 bp
* 28245 31504: contig of 3259 bp in length
* 31504 31604: gap of 100 bp
* 31604 35858: contig of 4255 bp in length
* 35859 35958: gap of 100 bp
* 35959 39610: contig of 3652 bp in length
* 39611 39710: gap of 100 bp
* 39711 42759: contig of 3049 bp in length
* 42760 42859: gap of 100 bp
* 42860 48014: contig of 5155 bp in length
* 48015 48114: gap of 100 bp
* 48115 52263: contig of 4149 bp in length
* 52264 52363: gap of 100 bp
* 52364 56374: contig of 4011 bp in length
* 56375 56474: gap of 100 bp
* 56475 63102: contig of 6628 bp in length
* 63103 63202: gap of 100 bp
* 63203 72687: contig of 9485 bp in length
* 72688 72787: gap of 100 bp
* 72788 97799: contig of 25012 bp in length
* 97800 97899: gap of 100 bp
* 97900 134325: contig of 36426 bp in length
* 134326 134425: gap of 100 bp
* 134426 167509: contig of 33084 bp in length.

FEATURES

source
1. .167509
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-45K21"
/clone_lib="RP11-45K21 Human Male BAC"
1. .1061
/note="assembly_fragment"
1162. .2381
/note="assembly_fragment"
2482. .3771
/note="assembly_fragment"
3872. .4949
/note="assembly_fragment"
5050. .6119
/note="assembly_fragment"
6220. .8023
/note="assembly_fragment"
8124. .10555
/note="assembly_fragment"
10656. .12768
/note="assembly_fragment"
12869. .15722
/note="assembly_fragment"
15823. .17959
/note="assembly_fragment"
18060. .21771
/note="assembly_fragment"
21872. .24539
/note="assembly_fragment"
24640. .28144
/note="assembly_fragment"
28245. .31503
/note="assembly_fragment"
31604. .35858
/note="assembly_fragment"
35959. .39610

misc_feature
/note="assembly_fragment"
39711. .42759
/note="assembly_fragment"
42860. .48014
/note="assembly_fragment"
48115. .52263
/note="assembly_fragment"
52364. .56374
/note="assembly_fragment"
56475. .63102
/note="assembly_fragment"
63203. .72687
/note="assembly_fragment"
clone_end:T7
vector_side:right"
72788. .97799
/note="assembly_fragment"
97900. .134325
/note="assembly_fragment"
134426. .167509
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

ORIGIN

Alignment Scores:
Pred. No.: 5,17e-172 Length: 167509
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.57% Indels: 0
DB: 2 Gaps: 0

US-10-066-521-6 (1-1344) x AC012107 (1-167509)

Qy 135 GlyHisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 154
|||
Db 47478 GGACATGGAGGTGACATCGGAGTACAGAGTCAAGTGTGATGACCAAAATTCGCTGAGGAG 47537

Qy 155 GluAspValArgArgSerPheGluAenThrAlaAlaAspTrpProGluMetGlnThrLeu 174
|||
Db 47538 GAGGATGTAGTGTAGTATTTTGAACACATGCTGCTGACTGGCCGGAATGCAAAACGTTG 47597

Qy 175 AlaGlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGly 194
|||
Db 47598 GCTGTGTCTTTTTCATTCAGACCGGTGGGCTTCGGGCTCGCAGGTGGTCTTCACCGGA 47657

Qy 195 LysSerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuGlyCysTrpAlaGlnGly 214
|||
Db 47658 AAGTCAGGAATGGGAAATCGGCTCTAGCCAGAAAGATCGTGTCTGTCTGGGGCAAGGT 47717

Qy 215 GlyLeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArg 234
|||
Db 47718 GGACTCTACCCAGGGAATGTTCTCCTACGTCTTCTCTCCCGTTAGAGAGATGCAGCGG 47777

Qy 235 LysLysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaPro 254
|||
Db 47778 AAGAAGGAGAGCAGTGTCTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCG 47837

Qy 255 ValThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAsp 274
|||
Db 47838 GTGACGAGATCATGTCCCGACCAGAAAGGCTGTTGTTTCATCATTCATTCGCGGTTTCGATGAC 47897

Qy 275 LeuGlySerValLeuAenAspThrLysLeuGlyLysAspTrpAlaGluLysGlnPro 294
|||
Db 47898 CTGGGCTCTGTCTCTCAACAATGACAAAAGCTCTGCAAAAGACTGGGCTGAGAGAGCAGCCT 47957

Qy 295 ProPheThrLeuIleArgSerLeuLeu 303
|||
Db 47958 CCGTTACCCCTCATACGAGTCTGCTG 47984

Search completed: July 16, 2004, 13:35:52

Job time : 12763 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 16, 2004, 06:23:54 ; Search time 1107 Seconds
(without alignments)
5157.704 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTFSYGLQWCLYE.....DDHSGVSWSLGAAGLEGLVS 1344

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10066521/runat_13072004_122316_11526/app_query.fasta_1.1543
-DB=N Geneseq_29Jan04 -QMT=fastap -SUFFIX=Oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10066521@cgn 1.1.819 @runat_13072004_122316_11526 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	4035	6 AAL44356	Aal44356 Human PYR
2	1006	74.9	3489	6 AAD41224	Aad41224 Human EMB
3	1006	74.9	3926	7 ADA45218	Ada45218 Human MAT
4	988	71.5	3830	7 ADA45220	Ada45220 Human MAT
5	968	72.0	3226	6 ABX97181	Abx97181 Human NOV
6	968	72.0	5859	6 AAL47135	Aal47135 Pyrin dom
7	968	72.0	6939	6 AAL47131	Aal47131 Pyrin dom
8	968	72.0	6939	6 AAL47140	Aal47140 Pyrin dom

9	616	45.8	3900	6	ABK48628	Abk48628 Human MAT
10	616	45.8	3900	7	AAD49018	Aad49018 Human MAT
11	385	28.6	1157	6	ABK48609	Abk48609 Human MAT
12	385	28.6	1157	7	AAD48999	Aad48999 Human MAT
13	196	14.6	1075	6	ABK48610	Abk48610 Human MAT
14	196	14.6	1075	7	AAD49000	Aad49000 Human MAT
15	12	0.9	3108	6	ADE36416	Ade36416 Human PAA
16	12	0.9	3186	6	AAL44363	Aal44363 Human PYR
17	12	0.9	3300	6	AAL47129	Aal47129 Pyrin dom
18	12	0.9	3306	9	ADC30316	Adc30316 Human nov
19	12	0.9	3447	6	ABK48611	Abk48611 Mouse MAT
20	12	0.9	3447	7	AAD49001	Aad49001 Mouse MAT
21	12	0.9	4931	6	ABL59333	AbL59333 Nucleotid
22	10	0.7	1722	9	ADE62511	Ade62511 Human gen
23	10	0.7	1722	9	ADE62507	Ade62507 Human gen
24	9	0.7	438	5	ABV38053	Abv38053 Human pro
25	9	0.7	578	4	AAL01744	Aal01744 Human rep
26	9	0.7	578	4	ABL97037	AbL97037 Human tes
27	9	0.7	693	6	ABQ40410	Abq40410 Oligonucl
28	9	0.7	693	6	ABQ40524	Abq40524 Oligonucl
29	9	0.7	693	6	ABQ40525	Abq40525 Oligonucl
30	9	0.7	693	6	ABQ40411	Abq40411 Oligonucl
31	9	0.7	702	4	AAK74564	Aak74564 Human imm
32	9	0.7	738	4	AAK59124	Aak59124 Human imm
33	9	0.7	2596	2	AAZ24899	Aaz24899 Human sec
34	9	0.7	4194	5	AAD02761	Aad02761 Human NB-
35	9	0.7	4329	5	AAD02762	Aad02762 Human NB-
36	9	0.7	4422	5	AAD02760	Aad02760 Human NB-
37	9	0.7	4466	5	AAD02765	Aad02765 Human NAC
38	9	0.7	4556	5	AAD02764	Aad02764 Human NAC
39	9	0.7	5100	7	ACC45152	Acc45152 Human cas
40	9	0.7	5122	4	AAD12951	Aad12951 Human G-p
41	9	0.7	5444	7	ACC45143	Acc45143 Human NAC
42	9	0.7	5477	2	AAT33514	Aat33514 Human DEC
43	9	0.7	5523	5	AAF83651	Aaf83651 Human CAR
44	9	0.7	5523	7	ABE55497	Abe55497 cDNA enco
45	9	0.7	5523	7	ABS56030	Abs56030 cDNA enco

ALIGNMENTS

RESULT 1
AAL44356
ID AAL44356 standard; cDNA; 4035 BP.

XX
AC AAL44356;
XX
DT 31-OCT-2002 (first entry)
XX
DE Human PYRIN-5 CDNA sequence.

XX Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-5.

XX Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 1..4035
FT /*tag= a
FT /product= "Human PYRIN-5"

XX WO200261049-A2.

XX 08-AUG-2002.

XX 31-JAN-2002; 2002WO-US0002967.

XX 31-JAN-2001; 2001US-0265231P.

PR 10-SEP-2001; 2001US-0318645P.
XX (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX Bertin J, Wang W, Blatcher M;
XX WPI; 2002-627477/67.
DR P-PSDB; AAO15585.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
PT diagnosing stress-related, apoptotic and inflammatory responses, or for
PT treating inflammatory and immune system disorders, cancers, or
PT neurological diseases.
XX
PS Claim 4; Fig 3; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
CC useful for modulating and diagnosing stress-related, apoptotic and
CC inflammatory responses. The PYRIN protein and DNA sequences are useful
CC for treating: inflammatory disorders and immune system disorders (e.g.
CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
CC (e.g. systemic lupus erythematosus and arthritis); and neurological
CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
CC protein and DNA sequences may also be used in screening assays, detection
CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
CC predictive medicine (e.g. diagnostic assays, clinical trials and
CC pharmacogenomics) and transcription profiling. The present DNA sequence
CC encodes the human PYRIN-5 protein
XX

SQ Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4035
Score: 1344.00 Matches: 1344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAL44356 (1-4035)

Qy 1 MetGluGlyAspLysSerLeuThrPheSerSerTyrGlyLeuGlnTrpCysLeuTyrGlu 20
Db 1 ATGGAAGGAGACAAATCGCTCACCTTTCCAGCTACGGGCTGCAATGGTGTCTCTATGAG 60
Qy 21 LeuAspLysGluGluPheGlnThrPheLysGluLeuLeuLysLysSerSerGluSer 40
Db 61 CTAGCAAGGAAGAATTTTCAGCAATTCAGGAAATTTACTAAAGAAGAATACTTCAGAATCG 120
Qy 41 ThrThrCysSerIleProGlnPheGluIleGluAsnAlaAsnValGluCysLeuAlaLeu 60
Db 121 ACCACATGCTCTATTCACAGTTTGAATCGAGATGCCACGTGGAAATGCTGGCACTC 180
Qy 61 LeuLeuHisGluTyrTyrGlyAlaSerLeuAlaTrpAlaThrSerIleSerIlePheGlu 80
Db 181 CTCTTGATGATGATTATGAGACATCGTGGCCCTGGGCTACGTCCATTAGCATCTTTGAA 240
Qy 81 AsnMetAsnLeuArgThrLeuSerGluLysAlaArgAspMetLysLysIleSerGln 100
Db 241 AACATGAACCTCGCAACCTCTCGGAAGGACACGGGATGACATGAAAAAATTTTCAA 300
Qy 101 AlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIleSerGlnAla 120
Db 301 GCTATGGACAAAGAGTGGCCACAGCAGCAGACAGACAGACAGACAAAGAAATTTCAAGCT 360
Qy 121 MetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGlyGlyAspThr 140
Db 361 ATGGAACAAAGAAGTGGCCACAGCAGCAGACAGACAGACAAAGACATGGAGGTGACACA 420

Qy 141 TrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspValArgArgSer 160
Db 421 TGGACCTACAAAGAGTCACGTGATGACAAATTCGCTGAGGAGGAGGATGACCTCGTAGT 480
Qy 161 PheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAlaPheAspSer 180
Db 481 TTTGAAAAACACTGCTGCTGACTGCGCCGGAATCAAAACGTTGGCTGGTCTTTGATTCA 540
Qy 181 AspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGlyIleGlyLys 200
Db 541 GACCGGTGGGGCTTCGGGCTCGCACGCGTGTCTGCGCAAGAAAGTCAGAAATGGGAAA 600
Qy 201 SerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnClyGlyLeuTyrGlnGlyMet 220
Db 601 TCGGCTCTAGCCAGAAAGGATCGTGTGCTGGCGGCAAGGTGGACTCTACCAAGGGAATG 660
Qy 221 PheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGluSerSerVal 240
Db 661 TTTCTCTACGTCTTCTTCTCCCGTTAGAGAGATGACGCGGAAGAGAGACAGTGTCT 720
Qy 241 ThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGluIleMetSer 260
Db 721 ACAGAGTTTCACTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGAGATCATGTCC 780
Qy 261 ArgProGluArgLeuPheIleLeuAspGlyPheAspAspLeuGlySerValLeuAsn 280
Db 781 CGACCAAGAAAGGCTGTGTTCATTCATGACGGTTTCATGACCTGGGCTCTGCTCTCAAC 840
Qy 281 AsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThrLeuIleArg 300
Db 841 AATGACACAAAGCTCTGCAAAAGATGGCTGAGAGACGCTCCGTTTCACTCATACGC 900
Qy 301 SerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrValArgAspVal 320
Db 901 AGTCTGCTGAGGAAGGTCTGCTGCTCCCTGATGCTCTGATGCTCCAGTCAGACGTG 960
Qy 321 GlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuValArgGlyIle 340
Db 961 GGCACAGAGAAAGCTCAAGTCAGAGGTGCTGCTCCCGCTTACCTGTGTAGTAGGAATC 1020
Qy 341 SerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHisGlnLysThr 360
Db 1021 TCCGGGGAACAAAGAAATCCACTTGTCTTTCAGCGCGGATGCTGAGCATCAGAAGACA 1080
Qy 361 GlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGlnValProAla 380
Db 1081 CAAGGGTTCGCTGCGATCATGAACAAACCGTGAAGTCTGCGACAGTCCAGGTCGCCGCC 1140
Qy 381 ValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGluSerValAla 400
Db 1141 GTGGGCTCTCTCATCTGCTGGCCCTGACAGCTGACAGACGTGGTGGGGAGAGCGTCGCC 1200
Qy 401 ProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGlnLeuThrPro 420
Db 1201 CCCTTCAACCAACCGCTCACAGGCTGACCGCGCTTTTGTGTTCATCAGCTCACCCCT 1260
Qy 421 ArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLysArgPheCys 440
Db 1261 CGAGGGGTGCTGCGGCGCTGTCTCAATCTGGAGAAAGAGTTGTCTCTGAAGCGCTTCTGC 1320
Qy 441 ArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAspLeuMet 460
Db 1321 CGTATGGCTGTGGAGGAGTGTGGAATAGAACTCAGTGTGTGATGGTACGACCTCATG 1380
Qy 461 ValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIleLeuLeuPro 480
Db 1381 GTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCAATGAACATCTTCTCCCA 1440
Qy 481 AspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAspPheCysAla 500
Db 1441 GACAGCCACTGTGAGGAGTACTACACTTCTTCCACTCTCAGTCTCCAGGACTTCTGTGCC 1500
Qy 501 AlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysProLeuTyrVal 520

Db 1501 GCCTTGTAAGTCTAGAGGCGCTGGAATCGAGCGAGCTCTCTGCGCTCTGTACGTT 1560
Qy 521 GluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHisSerLeuTrp 540
Db 1561 GAGAGACAAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCCATATCCATCCGCTTTGG 1620
Qy 541 MetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeuGluValLeu 560
Db 1621 ATGAAGCGCTTCTGTTGGCGCTCGTGAGCGAAGACGTAAGAGGCGCACCTGGAGGTCTG 1680
Qy 561 LeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpValSerLeuLeu 580
Db 1681 CTGGGCTGCTCCCGTTCCTCGGGTGAGAGGAGAGCTTCTGCACTGGGTCTCTCTGTTG 1740
Qy 581 GlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHisCysLeuPhe 600
Db 1741 GGTGAGAGCCTTAATGCCACCCAGGAGACACCTTGGAGCGCTTCCATCTGCTTTTC 1800
Qy 601 GluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGluValTrpLeu 620
Db 1801 GAGACTCAAGACAAAGAGTTTGTTCGCTTGGCAATTAACACACTTCCAAAGAGTCTGGCTT 1860
Qy 621 ProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHisCysProTyr 640
Db 1861 CCGATTAAACCAAGACCTGGACTTGATAGCATCTTCTTCTGCTCCAGCACTGTCCGTAT 1920
Qy 641 LeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSerAlaGluAla 660
Db 1921 TTGCGGAAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCA 1980
Qy 661 CysProValValProLeuTrpMetArgAspLysThrLeuIleGluGlnTrpGluAsp 680
Db 1981 TGTCTGTGGTCCCTCTATGATGCGGGATTAAGACCTCATTTGAGGAGCAGTGGGAAGAT 2040
Qy 681 PheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGlySerSerIle 700
Db 2041 TTCTGCTCCATGCTTGGCACCCACCAACCTGGCGGAGCTGGACCTGGGCGAGCATC 2100
Qy 701 LeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThrCysLysIle 720
Db 2101 CTGACAGCGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCTGCAAGATA 2160
Qy 721 GlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeuTrpArgIle 740
Db 2161 CAGACCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGCAGACCTCTGGAGATC 2220
Qy 741 ValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeuLysGluGlu 760
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCACCCACCTGAAGGAAG 2280
Qy 761 AspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGluSerLeuArg 780
Db 2281 GATGTAAGGATGGCGTGGAAGCCTTAAACACCCCAAAATGTTTGTGGAGTCTTTGAGG 2340
Qy 781 LeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIleLeuThrThr 800
Db 2341 CTGANTGCTGTGANTTGACCCATGCTGTTACTTGAAGATCTCCCAATCTCTTACGACC 2400
Qy 801 SerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGlnGlyValMet 820
Db 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGGTGACAGACCAGGAGTAATG 2460
Qy 821 ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIleLeuGluAsp 840
Db 2461 CCTCTAGTATGCTTGAGAGTCTCCAGGTGCGCCCTGCAAGAGCTGATCTGGAGGAC 2520
Qy 841 CysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSerAsnArgSer 860
Db 2521 TGTGGCATCACAGCACGGGTGCGAGAGTCTGGCCTCAGCCCTCAGCAACCGGAGC 2580
Qy 861 LeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsnLeuLeuCys 880

Db 2581 TTGACACACCTGTGCTCTATCAACAAACAGACCTGGGGAACGAAGGTGTAAATCTACTGTGT 2640
Qy 881 ArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGlnCysHisLeu 900
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCTAGTGCACCTG 2700
Qy 901 AspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrpLeuThrHis 920
Db 2701 GACACGGCTGGCTGTGTTTTCTTGCACCTTGCCTTATGGGTAACTCATGGCTCAGCAC 2760
Qy 921 LeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCysGluValMet 940
Db 2761 CTGAGCCTTAGCATGAACCCCTGTGGAACAATGGCGTGAAGCTTCTGTGCGAGGTCTATG 2820
Qy 941 ArgGluProSerCysHisLeuGlnAspLeuLysValLysCysHisLeuThrAlaAla 960
Db 2821 AGAAGAACCATCTTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTCACCGCGCG 2880
Qy 961 CysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSerLeuAspLeu 980
Db 2881 TGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGCCTGGATCTC 2940
Qy 981 ThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeuLysGlnLys 1000
Db 2941 ACGGACAAATCCCTGGGTGACGGTGGGGTTGCTGCGCTGTGCGAGGACTGAAGCAAAAG 3000
Qy 1001 AsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAspCysCysGlu 1020
Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTCTGATTTGCTGTGAG 3060
Qy 1021 AlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeuValGlnAsn 1040
Db 3061 GCATCTCTCTGGCCCTTCTCTGCACCGGCATCTGACCACTCTAAACCTGGTGCAGAT 3120
Qy 1041 AsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysProThrSerAsn 1060
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTTGCGCCTTTCCTGTCACCGCTCTAAC 3180
Qy 1061 LeuGlnIleIleGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLysLeuGlu 1080
Db 3181 TTACAGATAATTGGGCTGTGGAATGGCAGTACCTGTGCAAAATAAGGAAGCTGCTGGAG 3240
Qy 1081 GluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSerPheAspGlu 1100
Db 3241 GAAGTCAGCTACTCAAGCCCGCAGTCGTAATTCGCGGTAGTTGGCATTCTTTTGATGAA 3300
Qy 1101 AspAspArgHisLysIleGlyLeuThrPheArgLeuProGluSerArgAlaTrpProCys 1120
Db 3301 GATGACCGACACAAAATAGGACTTACTTTCGGCTCCCTGAAAGCCGGCATGGCCATGT 3360
Qy 1121 AlaLeuLeuTrpGlyMetAsnProGluGlnLysArgValSerLeuLeuAlaGlyAsp 1140
Db 3361 GCCTTGTGTGGGGATGAACCCAGAGCAGAAAGCGTGTGCTGCTTCTGGCTGGAGAC 3420
Qy 1141 PheLysSerSerThrArgPheAlaLysSerLeuCysLeuAlaThrAlaAsnGlyGluSer 1160
Db 3421 TTCAAGAGCAGTACACGATTTGGCAAGTCTCTGCTGGCCACGGCAATTTGTGAGTCC 3480
Qy 1161 GlnArgValAspAsnValGluGlnSerSerProGlnProMetAlaGlyThrGluHisLys 1180
Db 3481 CAGAGAGTTGACAAACGTGGAGCAGAGCTCCCGCAACCCCATGGCAGCAGCAACACAAA 3540
Qy 1181 GlnAspLysMetLeuSerValGlyTyrSerGlyValaTrpSerGluThrAlaGluLeuGlu 1200
Db 3541 CAAGATAAAATGTTGATGTTGGATATTCGGAGCCTGGTCTGAAACTGCTGAGCTCGAA 3600
Qy 1201 GlyLeuGlySerAsnSerAlaAspHisAspHisGlyGlyMetAlaTrpSerLeuGlyArg 1220
Db 3601 GGCGTTGGATCCCAACAGTGTGATCATGACCAAGGAGTATGGCCTGCTCCTAGGAGAG 3660
Qy 1221 GluLeuSerSerArgGlyLeuCysProThrValLeuMetThrThrAlaValCysProGly 1240
Db 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAACAGTGTGATGATGACCAAGCGGTGTGTCTCTGT 3720

QY 1241 HistTrpGluArgLeuGlySerArgGlyTrpCysLeuAsnSerAlaAspAspHisSerGly 1260
Db 3721 CACTGGGAGCGCTCTAGGGCTGGTGTCTTAACAGTGTGTATGACCAAGCGGT 3780
QY 1261 ValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSerAlaAspAsp 1280
Db 3781 GTGTCTGTGTCTACCTGGGAGCGCTGGCTCGAGGGCTTGTCTCAACAGTGTGTATGAC 3840
QY 1281 HisSerGlyValAlaAlaTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuValSerAsnSer 1300
Db 3841 CACAGCGTGTGTGTCTACTGGAGCGCTGGCTCGAGGGCTTGTGTCTCAACAGT 3900
QY 1301 AlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGluGlyLeuVal 1320
Db 3901 GCTGATGACCAACAGCGGTGTCTGTGTCTACTGGAGCGCTGGCTCGAGGGCTTGTGT 3960
QY 1321 SerAsnSerAlaAspAspHisSerGlyValSerTrpSerLeuGlyAlaAlaGlyLeuGlu 1340
Db 3961 TCCACAGTGTGTGTATGACCAACAGCGGTGTGTCTGTGTCTACTGGAGCGCTGGCTCGAG 4020
QY 1341 GlyLeuValSer 1344
Db 4021 GGGCTGTGTCT 4032
RESULT 2
AAD41224
ID AAD41224 standard; cDNA; 3489 BP.
XX AC AAD41224;
XX 30-OCT-2002 (first entry)
XX Human EMBRY-1 cDNA.
KW Human; embryogenesis associated protein; AIDS; reproductive disorder;
KW infertility; endometriosis; endometrial tumour; inflammatory disorder;
KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;
KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
KW EMBRY-1; allergy; gene therapy; gene; ss.
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 1..3489
XX FT /*tag= a
XX FT /product= "EMBRY-1 protein"
XX
XX PN WO200248362-A2.
XX PD 20-JUN-2002.
XX 14-NOV-2001; 2001WO-US043956.
XX PF 15-NOV-2000; 2000US-0249407P.
XX PR (INCY-) INCYTE GENOMICS INC.
XX FA
XX PI Ramkumar J, Arvizu C;
XX WPI; 2002-537629/57.
XX DR P-FSDB; AAE25053.
XX
XX New polypeptides of human embryogenesis associated proteins for screening
XX modulators useful for treating or preventing disorders e.g.
XX endometriosis, infertility, allergy, preeclampsia.
XX Claim 58; Page 95-96; 97pp; English.
XX
XX The invention relates to human embryogenesis associated proteins (EMBRY)
XX and nucleic acid molecules encoding such proteins. EMBRY sequences are
XX useful for screening modulators useful for treating or preventing
XX disorders associated with abnormal expression of EMBRY. The disorders

CC treated include reproductive disorders such as infertility,
CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,
CC contact dermatitis; disorders of the placenta such as preeclampsia,
CC abruptio placentae etc. Sequences of the invention are also useful for
CC analysing a proteome of a tissue or a cell type. EMBRY proteins are
CC useful as immunogens for preparing antibodies. Polynucleotides of the
CC invention are useful for creating knockin humanised animals or transgenic
CC animals to model human diseases. They are also used in gene therapy. The
XX present sequence is human EMBRY-1 cDNA
XX

SQ Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3489
Score: 1006.00 Matches: 1006
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.85% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAD41224 (1-3489)

QY 98 ILeSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluGluLe 117
Db 454 ATTTCAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGAAAT 513
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
Db 514 TCACAAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGACAAGAACAAGCATGGA 573
QY 138 GlyAspThrTrpAspTyrIysSerHisValMetThrIysPheAlaGluGluGluAspVal 157
Db 574 GGTGACACATGGGACTACAAAGAGTCAAGTATCACCATAATTCGCTGAGGAGGAGTGA 633
QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 634 CGTCGTAGTTTTGAAACACTGCTGCTGACTGCCCGAATGCAACAGTTGGCTGGTCT 693
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyIysSerGly 197
Db 694 TTTGATTACAGCGGTGGGCTTCCGGCTTCGACGCTCGCACGGTGTCTGCACGAAAGTCAGGA 753
QY 198 ILeGlyIysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyLeuTyr 217
Db 754 ATTTGGAAATTCGGCTCTAGCCAGAGGATCGTCTGTCTGGGCGCAAGGTGACTCTAC 813
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgIysLysGlu 237
Db 814 CAGGGAATGTTCTCCACGTCGTCTTCTTCTCCCGTTAGAGAGATGACGCGGAAGAAGGAG 873
QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 874 AGCAGTGTACAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCCGAGGCTCCGGTACGGAG 933
QY 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db 934 ATCATGTCCCGACACAGAAAGGTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
QY 278 ValLeuAsnAspThrIysLeuCysLysAspTrpAlaGluLysGlnProPheThr 297
Db 994 GTCTCAACAATACACAAGCTCTGCAAGAGTGGGCTGAGNAGCAGCTCCGTTCCACC 1053
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 1054 CTCATACGCACTGTCTGAGGAAGGTCTGCTCTCCCTGAGTCTCTCTGATGTCACCGTC 1113
QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValIserProArgTyrLeuLeuVal 337
Db 1114 AGAGAGCTGGGCAACAGAAAGCTCAAGTCAGAGGTCTGTCTCTCCCGTTACCTGTTAGTT 1173
QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357

Db 1174 AGAGGAATCTCCGGGAAACAAAGATCCACTTGTCTCTTGAGCGCGGATTCGTGAGCAT 1233
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 1234 CAGAAGACACAAAGGTTTGGTGGATCAATGAACAAACCGTGAGTGTCTCGACAGTGGCCAG 1293
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 1294 GTGCCCGCGTGGGCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGAG 1353
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 1354 AGCGTGGCCCCCTTCAACCAACCGCTCACAGCCCTGCACGCGCTTTGTGTTTCATCAG 1413
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluAtqValValLeuLys 437
Db 1414 CTCACCCCTCAGAGGCTGGTCCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTGAAG 1473
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCGCGTATGGCTGTGGAGGAGTGTGAATAGGAAGTCAGTGTCTTTGACGGTGAC 1533
Qy 458 AspLeuMetValGlnGlyLeuGlyCysLeuSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGGTTCAAAGGACTCGGGAGTCTGAGCTCCGCTGTCTGTTTCACATGAACATC 1593
Qy 478 LeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 1594 CTTCTCCAGACAGCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1653
Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTCTGTGCGCGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCCCT 1713
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTACGTTGAAGACAAAGAGGTCCATGGAGCTTAAACAGGCAGGCTTCCATATCCAC 1773
Qy 538 SerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluLeuAspValArgArgProLeu 557
Db 1774 TCGCTTTGGATGAGCGTTTCTTTTGGCTCTGTGAGCGAAGACGTAAGAGGCGCACTG 1833
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGGTCTCTGCTGGCTGCTCCCTCCCTGGGGTGAAGCAGAGAGCTTCTGCACTGGGTC 1893
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTTGGTTCAGCAGCCTTAATGCCACCCACCCAGGAGACACCCCTGGACGCTTCCAC 1953
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGTTGGCATTTAAACAGCTTCCAAGAA 2013
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGGCTTCCGATTAACCAAGAACTGGACTTGTAGCATCTTCTTCTGCTCCAGCAC 2073
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCCGATTTTTCGGGAAATTCGGGTGGATGTCAAAAGGGATCTTCCCAAGAGATGAGTCC 2133
Qy 658 AlaGluAlaCysProValValProLeuTyrMetArgAspLysThrLeuIleGluGluGln 677
Db 2134 GCTGAGGCATGTCTGTGGTCCCTCTATGGATGGGGATAAGACCCCTATTGAGGAGCAG 2193
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTTCTGCTCCATGCTTGGCACCCACCCACACCTGCGGAGCTGGACCTGGC 2253
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTCTGACAGAGCGGGCCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCCACC 2313

Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAGATACAGACCCCTGTATGTTTAGAAATGCACAGATTAACCCCTGGTGTGCAGACCTC 2373
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGGAGAATCGTCATGGCCAAACCGTAACCTAAGATCCCTCACTTGGGAGGACCCACCTG 2433
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisIleProLysCysLeuLeuGlu 777
Db 2434 AAGGAAGAGGATGTAAAGATGGCGTGTGAAGCCCTTAAACACCCCAAAATGTTTGTGGAG 2493
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCTTTGAGGCTGGATTGCTGTGGATTGACCATGCCCTGTACCTTGAAGATCTCCCAATC 2553
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTACGACCTCCCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGACCAG 2613
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAGTAATGGCTCTCAGTGTATGCTTGGAGAGTCTCCAGTGGCGCTGCAGAGCTGATA 2673
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGGAGACTGTGGCATCACAGCCACGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2734 AACCGAGCTTGACACACCTGTGCTTATCCAAACAACAGCCTGGGGAACGAAGGTGTAAT 2793
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2794 CTACTGTGTGCATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGTAGTGAATCAG 2853
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2854 TGCACCTGGACAGCGTGGCTGTGGTTTTCTTGGCACTTGGCTTATGGGGTAATCAAG 2913
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 2914 CTGAGCACCTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
Qy 938 GluValMetArgGluProSerCysHisLeuGluAspLeuGluLeuValLysCysHisLeu 957
Db 2974 GAGGTATGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3034 ACCCGCGTGTCTGAGAGTCTGTCTGTGTGATCTCGAGGAGACACACCTGAAGAGC 3093
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3094 CTGGATCTCAGGACAATGCCCTGGGTGACGGTGGGTTGCTGCACCTGTGCGAGGACTG 3153
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3154 AAGCAAAAGAACAGTGTCTGACGAGACTCGGTTGAAGGCACTGTGAGCTGACTTCTGAT 3213
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 3214 TGTGTGAGGCATCTCTCTTGGCCCTTTCTCTCAACCGGATCTGACACAGCTTAAACCTG 3273
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTCAGATAAATTCAGTCCCAAGGAATGATGAAGCTGTGTTCGGCTTGGCTTGTCTCC 3333
Qy 1058 ThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
Db 3334 ACGTCTAACTTACAGATAATTTGGGCTGTGGAAATGGCAGTACCCTGTGTGCAATAAGAG 3393

QY 1078 LeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSer 1097
 |||||
 Db 3394 CTGCTGGAGGAGTGCACCTACTCAAGCCCCAGTCGTAATTGACGGTAGTGGCAATTCT 3453
 |||||
 QY 1098 PheAspGluAspArg 1103
 |||||
 Db 3454 TTTGATGAAGATGACCGG 3471
 |||||
 RESULT 3
 ADA45218
 ID ADA45218 standard; DNA; 3926 BP.
 XX
 AC ADA45218;
 XX
 DT 20-NOV-2003 (first entry)
 XX Human MATER DNA.
 DE
 XX human ; MATER; maternal antigen that embryos require; NTP-ase;
 KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
 KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
 KW ovarian dysfunction; autoimmune premature ovarian failure;
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1. 3489
 FT CDS /*tag= a
 FT /*product= "MATER"
 FT
 XX
 PN EP1285964-A2.
 XX
 XX 26-FEB-2003.
 PD
 XX 12-JUL-2002; 2002EP-00090246.
 XX
 PF 10-AUG-2001; 2001DE-01039874.
 PR
 XX (SCHD) SCHERING AG.
 PA
 PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
 XX
 XX WPI; 2003-302814/30.
 DR P-FSDB; ADA45219.
 DR
 XX
 PT New nucleic acid for human maternal antigen that embryos require protein,
 PT useful e.g. in diagnosis and treatment of female infertility.
 XX
 PS Claim 1; Page 9-11; 31pp; German.
 XX
 CC This invention describes a novel human MATER (maternal antigen that
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
 CC associated with apoptosis, a defect in MATER activity causes growth
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
 CC MATER gene is found predominantly in ovary, testis and placenta. The
 CC products of the invention have antiinfertility, contraceptive,
 CC antiinflammatory, immunosuppressive and gynaecological activity.
 CC Effectors of the MATER polypeptide are used to treat infertility.
 CC associated with endometriosis and also ovarian dysfunction, autoimmune
 CC premature ovarian failure, inflammation, autoimmune diseases and female
 CC infertility, and as contraceptives. Measurements of autoantibodies
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful
 CC for diagnosis of female infertility, particularly by detecting mutations
 CC in the MATER gene with a DNA chip. This sequence encodes the human MATER
 CC polypeptide described in the invention.
 XX
 SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3926
 Score: 1006.00 Matches: 1006
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 74.85% Indels: 0
 DB: 7 Gaps: 0
 US-10-066-521-6 (1-1344) x ADA45218 (1-3926)
 QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGluIle 117
 |||||
 Db 454 ATTTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGAGACAGAAGAACAAAGAAATT 513
 |||||
 QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlnGlyHisGly 137
 |||||
 Db 514 TCACAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAGAACAAAGACATGGA 573
 |||||
 QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspVal 157
 |||||
 Db 574 GGTGACACATGGGACTACAAGAGTCACGTGATCACCATAATTCCGTGAGGAGGAGGATGTA 633
 |||||
 QY 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
 |||||
 Db 634 CGTCGTAGTTTTCAAAACACACTGCTGCTGACTGCCCGGAAATGCAAAACGTTGGCTGGTCT 693
 |||||
 QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
 |||||
 Db 694 TTTGATTACAGCCGCTGGGGCTTCCGGCCTCCGACCGTGGTCTCTGCACGGAAAGTCAGGA 753
 |||||
 QY 198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
 |||||
 Db 754 ATTTGGAAATCGGCTCTAGCCAGAAGATCGTCTGTCTGCTGGCGCAAGGTGACTCTAC 813
 |||||
 QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
 |||||
 Db 814 CAGGGAAATGTTCTCTACCGTCTTCTCTCCCGTTAGAGAGATGCAGCGGAAGAGGAG 873
 |||||
 QY 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
 |||||
 Db 874 AGCAGTGTACAGATTCACTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTACGGAG 933
 |||||
 QY 258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspLeuGlySer 277
 |||||
 Db 934 ATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 993
 |||||
 QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
 |||||
 Db 994 GTCTCAACAATGACACAAGCTCTGCAAGACTGGGCTGAGAAGCAGCGCTCGTTCACC 1053
 |||||
 QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
 |||||
 Db 1054 CTCATACGCAGTCTGCTGAGGAAGTCTCTGCTCCCTGAGTCTCTCTGATCGTCAACGTC 1113
 |||||
 QY 318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
 |||||
 Db 1114 AGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTGCTGTCTCCCGTTACCTGTTAGTT 1173
 |||||
 QY 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyHis 357
 |||||
 Db 1174 AGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTCTTGGCGCGGATTTGTTAGCAT 1233
 |||||
 QY 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
 |||||
 Db 1234 CAGAAGACACAAGGTTGGCTGGATCATGAACAACCGTGAGTGTCTCGACCACTGCCAG 1293
 |||||
 QY 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
 |||||
 Db 1294 GTCCCGCGCTGGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGGAG 1353
 |||||
 QY 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
 |||||
 Db 1354 AGCGTCGCCCTTCAACCAACCGCTCACAGGCTGCACGCCCTTTTGTGTTTCATCAG 1413
 |||||
 QY 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
 |||||
 Db 1414 CTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCTGAG 1473
 |||||

```
Qy 438 ArgPheCysArgMetAlaValIcLuGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 1474 CGCTTCGCGGTATGGCTGTGAGGGAGTGTGGAAATAGGAAGTCAGTGTTCGACGGTGAC 1533
Qy 458 AspLeuMetValGlnGlnGlyLeuGlyGluSerGlnLeuArgAlaLeuPheHisMetAsnIle 477
Db 1534 GACCTCATGGTTCAAGGACTCGGGAGTCTGAGCTCGGTCTCTGTGTTTCACATGAACATC 1593
Qy 478 LeuLeuProAspSerHisCysGluGluTyrTrpThrPhePheHisLeuSerLeuGlnAsp 497
Db 1594 CTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCCAGGAC 1653
Qy 498 PheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 1654 TTCTGTGCGCGCTTGACTACTAGTGTAGAGGGCTTGGAAATCGAGCCAGCTCTCTGCGCT 1713
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 1714 CTGTACGTTGAGAAAGACAAAGAGGTCCATGGAGCTTAAACAGGCGAGCTTCCATATCCAC 1773
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgAtqProLeu 557
Db 1774 TCGCTTTGGATGAAGCGTTCTTGTGTCCTCTGTGAGCGAGTAAACAGGCGAGCTTCCATATCCAC 1833
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpVal 577
Db 1834 GAGTCTCTGCTGGCTGTGTCGCTTCCCTGGGGGTGAAGCAGAAGCTTCTGCACTGGGTC 1893
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 1894 TCTCTGTTGGGTTCAGCAGCCTTAATGCCACCAACCCAGGAGACACCTTGGAGCGCTTCCAC 1953
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 1954 TGTCTTTTCGAGACTCAAGACAAAGAGTTGTTGCTTGGCATTTAAACAGCTTCCAAAGAA 2013
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 2014 GTGTGGCTTCCGATTAACCAAGAACTTGACTTGTATAGCATCTTCTCTGCTCCAGCAC 2073
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 2074 TGTCCGATTTTCGGAATAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGAGTCC 2133
Qy 658 AlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluGlnGln 677
Db 2134 GCTGAGGCATGTCTGTGGTCCCTCTATGGATGCGGGATAAGACCCCTCATTTGAGGAGCAG 2193
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 2194 TGGGAAGATTTCTGCTCATGCTTGGCACCCACCAACCTGCGGAGCTGGACCTGGGC 2253
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 2254 AGCAGCATCTCTGACAGACGGCGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCAAC 2313
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 2314 TGCAGGATACAGACCCCTGATGTTTAGAAAATGCAAGATTAACCTGTGTGTGAGCACCTC 2373
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 2374 TGGAGATCTGTATGGCCACACCGTAACCTAGATCCCTCAACTTGGAGGAGCCACCACTG 2433
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 2434 AAGGAAGAGGATGTAAGGATGGCGTGTGAAGCCTTAAACACACCCCAAAATGTTTGTGAG 2493
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 2494 TCTTTGAGGCTGGATGCTGTGGATTGACCCATGCCCTGTTACCTGAAGATCTTCCCAATC 2553
```

```
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 2554 CTTTACGACCTCCCCAGCCTGAAATCTCTGAGCCTGCGAGAAACAAGGTGACAGACCAG 2613
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 2614 GGAAGTAATGCTCTCAGTGTATGCTTTGAGAGTCTCCAGTGCCTGCAAGAGCTGATA 2673
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 2674 CTGAGAGACTGTGGCATCACGCCACGGTTGCAGAGTCTGGCCTCAGCCCTCGTCAGC 2733
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 2734 AACCGGAGCTTGACACACCTGTGCTTATCCAAACAACAGCCTGGGGAACGAAGGTGAAAT 2793
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 2794 CTACTGTGTGATCCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 2853
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 2854 TGCACCTGACACGGCTGGCTGTGTTTCTTGCACCTTGGCTTATGGGTAACTCATGG 2913
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValIcLuAspAsnGlyValLysLeuLeuCys 937
Db 2914 CTGACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTTCTGTGC 2973
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuIleValLysCysHisLeu 957
Db 2974 GAGGTATGAGAAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 3033
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 3034 ACCGCGCGTGTGTGAGAGTCTGTCTGTGTGTATCTCGAGGAGCAGACACCTGAAGAGC 3093
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 3094 CTGGATCTCAGGACAAATGCCCTGGGTGACGGTGGGGTGTCTGCACCTGTGCGAGGAGCTG 3153
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 3154 AAGCAAAAGAAACAGTGTCTTGACGAGACTCGGGTTGAAGGCATGTGAGCTGACTTCTGAT 3213
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 3214 TGCTGTGAGGACATCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTTAAACCTG 3273
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 3274 GTGCAGAAATAACTTCACTCCCAAGGAATGATGAAGCTGTGTTCGGCTTTGGCTGTCCC 3333
Qy 1058 ThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIleArgLys 1077
Db 3334 ACCTCTAACTTACAGATAATTTGGGCTGTGGAATGGCAGTACCCTGTGCAATAAAGAGAG 3393
Qy 1078 LeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHisSer 1097
Db 3394 CTCTGGAGGAAGTGAGCTACTCAAGCCCGGAGTCTGTAATTGACGGTAGTGTGGCATCTC 3453
Qy 1098 PheAspGluAspArg 1103
Db 3454 TTTGATGAAGATGACCGG 3471
RESULT 4
ADA45220
ID ADA45220 standard; DNA; 3830 BP.
XX
AC ADA45220;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human MATER splice variant DNA SEQ ID 3.
```

XX human; MATER: maternal antigen that embryos require; NTP-ase;
 KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;
 KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
 KW ovarian dysfunction; autoimmune premature ovarian failure;
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3432
 FT /*tag= a
 FT /product= "splice variant of MATER"
 XX
 PN EP1285964-A2.
 XX
 PD 26-FEB-2003.
 XX
 PF 12-JUL-2002; 2002EP-00090246.
 XX
 PR 10-AUG-2001; 2001DE-01039874.
 XX
 PA (SCHD) SCHERING AG.
 XX
 XX Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
 PI
 XX WPI: 2003-302814/30.
 DR P-PSDB; ADA45221.
 XX
 PT New nucleic acid for human maternal antigen that embryos require protein,
 PT useful e.g. in diagnosis and treatment of female infertility.
 XX
 PS Claim 1; Page 17-18; 31pp; German.
 XX
 CC This invention describes a novel human MATER (maternal antigen that
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
 CC associated with apoptosis, a defect in MATER activity causes growth
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
 CC MATER gene is found predominantly in ovary, testis and placenta. The
 CC products of the invention have antiinfertility, contraceptive,
 CC antiinflammatory, immunosuppressive and gynaecological activity.
 CC Effectors of the MATER polypeptide are used to treat infertility.
 CC associated with endometriosis and also ovarian dysfunction, autoimmune
 CC premature ovarian failure, inflammation, autoimmune diseases and female
 CC infertility, and as contraceptives. Measurements of autoantibodies
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful
 CC for diagnosis of female infertility, particularly by detecting mutations
 CC in the MATER gene with a DNA chip. This sequence represents a splice
 CC variant of human MATER in which exon 4 of the usual form is absent.
 XX
 SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3830
 Score: 988.00 Matches: 988
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.51% Indels: 0
 DB: 7 Gaps: 0
 US-10-066-521-6 (1-1344) x ADA45220 (1-3830)
 Qy 116 GluileSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGluGlnGly 135
 Db 451 GAAATTTTCAACAGCTATGGACAAAGAGGTGCCACAGCAGCAGACAGAGAACAAGGA 510
 Qy 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGlu 155
 Db 511 CATGGAGGTGACATGGGACTACAGAGTCAGCTGATGACCAATTCGCTGAGGAGGAG 570
 Qy 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
 Db 571 GATGTACGTCTGCTAGTTTGTGAAACACACTGCTGCTGACTGGCGGGAATGCAACGTTGGCT 630

Qy 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
 Db 631 GGTGCTTTTATTTCACACCGGTGGGCTTCGGGCCCTCGCACGGTGGTTCTGCAAGGAAAG 690
 Qy 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGly 215
 Db 691 TCAGGAATTGGGAATCGGCTTAGCCACAGAGATCGTGTCTGTGCTGGGCGCAAGGTGA 750
 Qy 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
 Db 751 CTCTACAGGGAATGTTCTCTACGTCTTCTCTCTCCCGTTAGAGAGATGCAGCGGAAG 810
 Qy 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
 Db 811 AAGGAGCAGCAGTGTACAGAGTTTCATCTCAGGGAGTGGCCAGACTCCACAGGCTCGGGT 870
 Qy 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspLeu 275
 Db 871 ACGAGATCATGTCCCGACACAGAAAGCTGTGTTCATTCATTGACCGTTCGATGACCTG 930
 Qy 276 GlySerValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPro 295
 Db 931 GGCTCTGTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAAGCAGCCTCG 990
 Qy 296 PheThrIleIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
 Db 991 TTCACCTCATACGCAGTCTGTCTGAGGAAGGTCTCTCTCCCTGAGTCTCTTCCTGATCGT 1050
 Qy 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
 Db 1051 ACCGTCAGAGAGCTGGGCACAGAGAGCTCAAGTCAGAGTCTGTCTCTCCCTTACCTG 1110
 Qy 336 LeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGly 355
 Db 1111 TTAGTTAGAGGAATCTCCGGGGAACAAGAATCCACTTGTCTCTTCAGGCGCGGATGGT 1170
 Qy 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
 Db 1171 GAGCATCAAGAGACACAGGGTTGCGTGGGATCATGAACAACCGTAGCTGCTCGACCCAG 1230
 Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
 Db 1231 TGCAGGTGCCCGCTGGGCTCTCTCATCTGGTGGCCCTGCAGCTGCAGGACGTGGTG 1290
 Qy 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
 Db 1291 GGGGAGAGCGTCCGCCCTTCAACCAACGCTCACAGGCTGCACGCGCTTTTGTGTTT 1350
 Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValVal 435
 Db 1351 CATCAGTCAACCTCGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAGAGTTGTC 1410
 Qy 436 LeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
 Db 1411 CTGAAGCGCTTCGCCGTATGGCTGTGGAGGGAGTGTGGAAAGAGTACAGTGTGTGAC 1470
 Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMet 475
 Db 1471 GGTGACGACCTCATGTTTCAAGGACTCGGGGAGTCTGAGTCTCGTCTCTGTTTCATG 1530
 Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeu 495
 Db 1531 AACATCTTCTCCACAGACGCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTC 1590
 Qy 496 GlnAspPheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeu 515
 Db 1591 CAGGACTTCTGTGCCCTTGTACTACGTGTAGAGGGCTTGAATCGAGCCAGCTCTC 1650
 Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
 Db 1651 TGCCCTCTGTACGTTGAGAAGACAAAGAGTCCATGGAGTCTAAACAGGAGGCTTCCAT 1710

QY 536 IleHisSerLeuTrpMetIysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
DB 1711 ATCCACTCGCTTGGATTGAAGCGTTTCTTGTGGCCCTCGTGAGCGAAGACGTAAGGAGG 1770
QY 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
DB 1771 CCACGTGGAGGTCCTGCTGGGTGTCCCGTTCCCTTGGGGGTGAAGCAGAACCTTCTGCAC 1830
QY 576 TrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
DB 1831 TGGGTCTCTCTGTGGGTGAGCAGCGCTAATGCCACCCAGCAGACACCCCTGGACGCC 1890
QY 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
DB 1891 TTCACCTGCTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCCTTGGCATTAACACAGCTTC 1950
QY 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
DB 1951 CAGAGAGTGTGGCTTCCGATTAAACAGAACCTGGACTTGATAGCATCTTCTTCTGCCCTC 2010
QY 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
DB 2011 CAGCACGTCTCGTATTTTCGGGAAATTCGGGTGGATGTCAAGGGGATCTTCCCAAGAGAT 2070
QY 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
DB 2071 GAGTCCGCTGAGGAGTGTCTGTGTGTCCTCATGGATGCGGATAGACCCCTCATTTGAG 2130
QY 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
DB 2131 GAGCAGTGGGAAGATTTCTGCTCATGCTTGGCACCCACCCAGCCTGGCGCAGCTGGAC 2190
QY 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
DB 2191 CTGGGCACGACATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCAT 2250
QY 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
DB 2251 CCCACCTGCAAGATACAGACCTGTATGTTTGAAGATGCACAGATTACCCCTGTGTGCAG 2310
QY 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyThr 755
DB 2311 CACCTCTGGAGAATCGTATGGCCAAACCGTAAGATCCCTCAACTTGGGGAGGCACC 2370
QY 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
DB 2371 CACCTGAAGGAGAGGATGAAGATGGCGTGTGAGCCCTTAAACACCCCAAAATGTTTG 2430
QY 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSer 795
DB 2431 TTGGAGTCTTTGAGGCTGGATTGCTGTGATTGACCCATGCTGTACCTGAAGATCTCC 2490
QY 796 GlnIleLeuThrThrSerProSerIleLysSerLeuSerLeuAlaGlyAsnLysValThr 815
DB 2491 CAAATCCTTACGACCTCCCCAGCCTGAAATCTCTGAGCCCTGGCAGGAAACAAGGTGACA 2550
QY 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
DB 2551 GACCAGGGAGTAAATGCCTCTCAGTGTAGCTTCCAGATGCTCCAGTGGCCCTGCAGAG 2610
QY 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
DB 2611 CTGATCTGGAGGACTGTGGCATCACAGCCAGCGGTTGCCAGAGTCTTGGCCTCAGCCCTC 2670
QY 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
DB 2671 GTCAGCAACCCGAGCTTGACACACTGTGCCATATCCAACACACACCCCTGGGGAGCAAGGT 2730
QY 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
DB 2731 GTAATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTGTGACAGAGCTGATGCTG 2790
QY 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915

DB 2791 AATCAGTGCACCTGGACACGCTGGCTGTGGTTTCTTGCACCTTGGCTTATGGGTAAC 2850
QY 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
DB 2851 TCATGGCTGACGACCTTGAGCCCTTAGCATGAACCTGTGGAAGACAATGGCGTGAAGCTT 2910
QY 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
DB 2911 CTGTGCGAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGT 2970
QY 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValLysSerArgSerArgHisLeu 975
DB 2971 CATCTCACCCCGCGTCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTG 3030
QY 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGlu 995
DB 3031 AAGAGCCTGGATCTCACGACAATGCCCTGGGTGACGCTGGGGTTGCTGCACCTGTGCGAG 3090
QY 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
DB 3091 GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACT 3150
QY 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
DB 3151 TCTGATTGCTGTGAGGCACCTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCAGTCTA 3210
QY 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
DB 3211 AACCTGGTGAGAAATTAACCTTCACTGCAAGGAATGATGAAGCTGTGTCGGCCTTTGGCC 3270
QY 1056 CysProThrSerAsnLeuGlnIleGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
DB 3271 TGTCCCACGCTTAACCTTACAGATAATTGGGCTGTGGAATGGCAGTACCCTGTGCAATA 3330
QY 1076 ArgLysLeuLeuGluGluValGlnLeuLeuLysProArgValLysLeuAspGlySerTrp 1095
DB 3331 AGGAAGCTGTGGAAGGAGTGCAGACTACTCAAGCCCGAGTCTGTAATTGACGGTAGTTGG 3390
QY 1096 HisSerPheAspGluAspArg 1103
DB 3391 CATCTTTTGTGAAGATGACCCG 3414
RESULT 5
ABX97181
ID ABX97181 standard; cDNA; 3226 BP.
XX
AC ABX97181;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV125b cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P. *not ok*
PR 08-MAR-2001; 2001US-0274194P. *no*
PR 08-MAR-2001; 2001US-0274281P. *no*
PR 08-MAR-2001; 2001US-0274322P. *no*
PR 09-MAR-2001; 2001US-0274849P. *no*
PR 12-MAR-2001; 2001US-0275235P. *no*
PR 13-MAR-2001; 2001US-0275578P. *no*
PR 13-MAR-2001; 2001US-0275579P. *no*

13-MAR-2001; 2001US-0275601P. *10*
14-MAR-2001; 2001US-0276000P. *10*
16-MAR-2001; 2001US-027676P. *10*
19-MAR-2001; 2001US-0276994P. *10*
20-MAR-2001; 2001US-0277239P. *10*
20-MAR-2001; 2001US-0277321P. *10*
20-MAR-2001; 2001US-0277327P. *10*
21-MAR-2001; 2001US-0277791P. *10*
22-MAR-2001; 2001US-0277833P. *10*
23-MAR-2001; 2001US-0278152P. *10*
26-MAR-2001; 2001US-0278894P. *10*
27-MAR-2001; 2001US-0278999P. *10*
27-MAR-2001; 2001US-0279036P. *10*
28-MAR-2001; 2001US-0279344P. *10*
30-MAR-2001; 2001US-0279388P. *10*
30-MAR-2001; 2001US-0279959P. *10*
30-MAR-2001; 2001US-0280233P. *10*
02-APR-2001; 2001US-0280802P. *10*
02-APR-2001; 2001US-0280822P. *10*
02-APR-2001; 2001US-0280900P. *10*
04-APR-2001; 2001US-0281194P. *10*
13-APR-2001; 2001US-0283675P. *10*
30-APR-2001; 2001US-0287424P. *10*
03-MAY-2001; 2001US-0288066P. *10*
03-MAY-2001; 2001US-0288342P. *10*
15-MAY-2001; 2001US-0288528P. *10*
16-MAY-2001; 2001US-0291099P. *10*
16-MAY-2001; 2001US-0291240P. *10*
30-MAY-2001; 2001US-0294485P. *10*
31-MAY-2001; 2001US-0294889P. *10*
31-MAY-2001; 2001US-0294899P. *10*
18-JUN-2001; 2001US-0299027P. *10*
19-JUN-2001; 2001US-0299303P. *10*
19-JUN-2001; 2001US-0299310P. *10*
10-JUL-2001; 2001US-0304354P. *10*
31-JUL-2001; 2001US-0309198P. *10*
16-AUG-2001; 2001US-0312903P. *10*
10-SEP-2001; 2001US-0318482P. *10*
12-SEP-2001; 2001US-0318770P. *10*
27-SEP-2001; 2001US-0325430P. *10*
27-SEP-2001; 2001US-0325681P. *10*
18-OCT-2001; 2001US-0330380P. *10*
31-OCT-2001; 2001US-0335301P. *10*
14-NOV-2001; 2001US-0332172P. *10*
14-NOV-2001; 2001US-0332271P. *10*
14-NOV-2001; 2001US-0332272P. *10*
14-NOV-2001; 2001US-0333184P. *10*
14-NOV-2001; 2001US-0333272P. *10*
21-NOV-2001; 2001US-0332094P. *10*
03-DEC-2001; 2001US-0337426P. *10*
03-DEC-2001; 2001US-0338092P. *10*
04-DEC-2001; 2001US-0337185P. *10*
03-JAN-2002; 2002US-0345705P. *10*
07-MAR-2002; 2002US-00092900. *10*

(CURA-) CURAGEN CORP.

Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
Lepley DM, Rieger DK;
WPI; 2002-723332/78.
P-PSDB; ABU65214.

NOVX polypeptides and polynucleotides, useful for preventing or treating
a disorder associated with aberrant NOVX expression or activity e.g.,
cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
asthma.

PS Claim 13; Page 467; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
activity. Pharmaceutical compositions comprising the NOVX proteins or
nucleic acid molecules or NOVX antibodies are useful for preventing or
treating a disorder associated with aberrant NOVX expression or activity
e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
asthma. The products of the invention can be used for gene therapy or in
a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX

SQ Sequence 3226 BP; 738 A; 856 C; 914 G; 718 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3226
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x ABX97181 (1-3226)

QY 136 HisGlyGlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGlu 155
Db 249 CATGGAGGTGACATACATGGAGGTACAGAGTTCAGTATGACCAATTCGCTGAGGAGGAG 308
QY 156 AspValArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAla 175
Db 309 GATGTACGTCGTAGTTTGAACACTGCTGTGCTGCTGCGGAAATGCAACGTTGGCT 368
QY 176 GlyAlaPheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLys 195
Db 369 GGTGCTTTTGAATTCAGACCGGTGGGGCTTCGGGCTCGCACGGTGGTTCTGCACGGAAAG 428
QY 196 SerGlyIleGlyLysSerAlaLeuAlaArgArgGileValLeuLysTrpAlaGlnGly 215
Db 429 TCAGGAATTGGGAAATCGGCTCTAGCCAGAAAGATCGTGTCTGTCTGGCGCAAGGTGGA 488
QY 216 LeuTyrGlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLys 235
Db 489 CTCTACCGAGGAATGTTCTCTACGTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 548
QY 236 LysGluSerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProVal 255
Db 549 AAGGAGAGCAGTGTACAGAGTTCATCTCCAGGAGTGGCCAGACTCCCGAGGCTCCGGTG 608
QY 256 ThrGluIleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspLeu 275
Db 609 ACGGAGATCATGTCCCGACAGAAAGCTGTGTTTCATCATTTGACGGTTTCGATGACCTG 668
QY 276 GlySerValLeuAsnAspThrLysLeuLysLysAspTrpAlaGluLysGlnProPro 295
Db 669 GGCTCTGTCTCAACAATGACAAAGCTCTGCAAGAGCTGGGCTGAGAGAGCGCTCCG 728
QY 296 PheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleVal 315
Db 729 TTCACCTCATACGAGTCTGCTGAGGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 788
QY 316 ThrValArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeu 335
Db 789 ACCGTGAGAGCTGGGCGACAGAGAGCTCAAGTCAGAGTCTGTCTCTCTCTCTCTCTCTCT 848
QY 336 LeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGly 355
Db 849 TTAGTTAGAGGAATCTCCGGGGAACAAAGAACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 908
QY 356 GluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGln 375
Db 909 GAGCATCAGAGACACAGGGTTGGTGGATCATGAACACCGTGGAGTGTCTCTCTCTCTCTCTCT 968

1075
33226
Check these
on Monday

Qy 376 CysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValVal 395
Db TGCAGAGTGCCTGGCTCTCTCATCTGGTGGCCCTGCAGCTGCAGACGTGGTG 1028
Qy 396 GlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPhe 415
Db GGGGAGAGCGTGGCCCTTCAACCAACCGCTCAAGCCCTGCACGCGCTTTTGTGTTT 1088
Qy 416 HisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValVal 435
Db CATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGAGTTGTC 1148
Qy 436 LeuIysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAsp 455
Db CTGAAGCGCTTCTGCGGTATGGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTTCGAC 1208
Qy 456 GlyAspAspLeuMetValGlnGlyLeuGlyCysGluSerGlnLeuArgAlaLeuPheHisMet 475
Db GGTGACGACCTCATGGTTCAAGGACTCGGGAGTCTGAGCTCCGTGTCTGTTCACATG 1268
Qy 476 AsnIleLeuLeuProAspSerHisCysGluGluTyrTyrPhePheHisLeuSerLeu 495
Db AACATCCTTCTCCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAGTCTC 1328
Qy 496 GlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeu 515
Db CAGGACTTCTGCGCCCTTGTACTACGTGTAGAGGCTCGGAATCGAGCCAGCTCTC 1388
Qy 516 CysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHis 535
Db TGGCTCTCTGTGGTTCAGACGACAAAGAGTCCATGGAGCTTAAACAGGCGAGGCTCCAT 1448
Qy 536 IleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArg 555
Db ATCCACTCGCTTGGATGAACGGTTCTTGTGTGGCTCGAGGAGGAGCGTAAGGAGG 1508
Qy 556 ProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHis 575
Db CCAGTGGAGTCTGCTGGCTGCTCCCTTCCCTGGGGGTGAAGCAGAGCTTCTGCAC 1568
Qy 576 TrpValSerLeuLeuGlyGlnProAsnAlaThrThrProGlyAspThrLeuAspAla 595
Db TGGGTCTCTCTGTGGTTCAGACGCTAAATGCCACACCCCGAGGACACCCCTGGAGCC 1628
Qy 596 PheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPhe 615
Db TTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTC 1688
Qy 616 GlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeu 635
Db CAAGAAGTGTGGCTTCCGATTAACAGAACCTGGACTTGTATGATCATCTTCTTCTGCTC 1748
Qy 636 GlnHisCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAsp 655
Db CAGCACTGTCCGTATTTGGGAAATTCGGGTGGATGTCAAAAGGATCTTCCCAAGAGAT 1808
Qy 656 GluSerAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGlu 675
Db GAGTCCGCTGAGGCATGTCTGTGGTCCCTCATATGGATGCGGGATAAGACCTCATTTGAG 1868
Qy 676 GluGlnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAsp 695
Db GAGCAGTGGGAGATTTCTGCTCATCTGTGGCACCCACCCACACCTTCGCGCAGCTGGAC 1928
Qy 696 LeuGlySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHis 715
Db CTGGGCAGCAGCATCTCTGACAGAGCGGCGCCATGAAGACCTGTGTGTCGAAGCTGAGGCAT 1988
Qy 716 ProThrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGln 735
Db CCCACCTGCAAGATACAGACCCCTGATGTTAGAAATGCACAGATTACCCCTGGTGTGCAG 2048

Qy 736 HisLeuTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThr 755
Db CACCTCTGGAGATCGTCAATGGCCCAACCGTAACCTAAGATCCCTCACTTGGGAGGCACC 2108
Qy 756 HisLeuLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeu 775
Db CACCTGAAGGAAGAGATGTAAAGATGGCTGTGAAGCTTAAACACACCCCAAAATGTTTG 2168
Qy 776 LeuGluSerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuIleSer 795
Db TTGAGTCTTTGAGGCTGGATGCTGTGGATGACCCATGCCTGTACCTGGAAGATCTCC 2228
Qy 796 GlnIleLeuThrThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThr 815
Db CAAATCCTTAGGACCTCCCCCAGCCTGAATCTCTGAGCCTGGCAGGAACACAGGTGACA 2288
Qy 816 AspGlnGlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLys 835
Db GACCAGGAGTAATGCCCTCTCAGTGATGCCCTTGAGAGTCTCCAGTGCGCCCTGCAGAAG 2348
Qy 836 LeuIleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeu 855
Db CTGATCTGGAGGACTGTGGCATCACAGCCACGGGTGCCAGAGTCTGGCCTCAGCCCTC 2408
Qy 856 ValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGly 875
Db GTCAGCAACCGGAGCTTGACACACTGTGCTATCCAAACACAGCCTGGGGACAGAGGT 2468
Qy 876 ValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeu 895
Db GTAAATCTACTGTGTGCATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTG 2528
Qy 896 AsnGlnCysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsn 915
Db AATCAGTGCACCTGACACGCGCTGCTGTGGTGTTCCTTGCACTTGGCTTATGGGTAAAC 2588
Qy 916 SerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeu 935
Db TCATGCTGACGACCTGAGCCTTAGCATGAAACCTCTGGAGAACAAATGGCGTGAAGCTT 2648
Qy 936 LeuCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCys 955
Db CTGTGCGAGTCAATGAGAGAACCATCTGTGTCATCTCCAGGACCTGGAGTTGGTAAAGTGT 2708
Qy 956 HisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeu 975
Db CATCTCACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTG 2768
Qy 976 LysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGlu 995
Db AAGAGCTGGATCTCACGGACAAATGCTGGGTGACGGTGGGGTTCGTGCACTGTGCGAG 2828
Qy 996 GlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThr 1015
Db GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATCTGGACTGACT 2888
Qy 1016 SerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeu 1035
Db TCTGATGTGTGAGGACCTCTCTTGGCCCTTCTCTGCAACCGGCACTCTGACACAGTCTA 2948
Qy 1036 AsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAla 1055
Db AACCTGGTCAGANTAACTTCAGTCCCAAGGAANTGATGAAGCTGTGTTTCGGCTTTGCC 3008
Qy 1056 CysProThrSerAsnLeuGlnIleLeuGlyLeuTrpLysTrpGlnTyrProValGlnIle 1075
Db TGTCCCACGCTCTAACTTACAGATAATTTGGCTGTGGAAATGGCAGTACCTCTGCAATA 3068
Qy 1076 ArgLysLeuLeuGluValGlnLeuLeuLysProArgValValIleAspGlySerTrp 1095
Db AGGAAGCTGTGGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTTGACGCTAGTTGG 3128
Qy 1096 HisSerPheAspGluAspArgArg 1103

```

Db      3129 CATTCTTTTGTGATGAAGATGACCGG 3152
|||||
RESULT 6
AAL47135
ID      AAL47135 standard; DNA; 5859 BP.
XX
AC      AAL47135;
XX
DT      20-AUG-2002 (first entry)
XX
XX      Pyrin domain containing protein NALP8/Py12 coding sequence.
DE
KW      Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW      antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW      neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW      nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW      inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW      arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW      osteoarthritis; glomerulonephritis; gene; ds.
XX
OS      Unidentified.
XX
PN      WO200240668-A2.
XX
PD      23-MAY-2002.
XX
PF      30-OCT-2001; 2001WO-EP012545.
XX
PR      15-NOV-2000; 2000DE-01056687.
XX
PR      30-NOV-2000; 2000DE-01059595.
XX
PA      (APOT-) APOTECH RES & DEV LTD.
XX
PI      Tschoep J, Martinon F;
XX
DR      WPI; 2002-427093/45.
XX
DR      P-PSDB; RAO17863.
XX
PT      New DNA encoding protein with pyrin domain, useful for treating diseases
PT      involving impaired signal transduction, particularly inflammation, also
PT      proteins and antibodies.
XX
PS      Claim 5; Fig 1; 116pp; German.
XX
CC      The present invention relates the DNA and their encoded proteins, where
CC      the proteins contain at least one PYD (pyrin) domain. These can be used
CC      to treat diseases associated with impaired intracellular signal
CC      transduction, particularly inflammation such as psoriasis,
CC      arteriosclerosis, bacterial or viral infections (particularly meningitis
CC      and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC      sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC      and Parkinson's diseases. The present sequence is a coding sequence of
CC      the invention
XX
SQ      Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      0      Length:      5859
Score:          968.00  Matches:      968
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      72.02%      Indels:      0
DB:              Gaps:      0

US-10-066-521-6 (1-1344) x AAL47135 (1-5859)

Qy      98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluLeu 117
Db      2389 ATTTCAAGCTATGGAAACAGAGAGGTGCCACAGCAGACAGACAGAAACAGAAATT 2448
Qy      118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
|||||

```

```

Db      2449 TCACAAAGCTATGGAAACAGAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGACATGGA 2508
Qy      138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db      2509 GGTGACACATGGAGCTACAGAGAGTCAGTGATGACCAAAATTCGCTGAGGAGGAGATGA 2568
Qy      158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db      2569 CGTCGTAGTTTTGAAAACACTGCTGCTGACTGGCGGGAATGCAAAAGTTGGCTGGTGCT 2628
Qy      178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db      2629 TTTGATTACAGCCGTGGGGCTTCGGCCTCGCACGGTGGTTCGACGCAAAAGTCAGGA 2688
Qy      198 IleGlyLysSerAlaLeuAlaArgAlaGlnIleValLeuCysTrpAlaGlnGlyLeuTyr 217
Db      2689 ATTGGGAAATCGGCTTACGCAAGAGGATCGTGCTGTGTGGCGCAAGGTGGACTCTAC 2748
Qy      218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db      2749 CAGGGAAATGTTCTCTACGCTCTTCTTCTCCCGTTAGAGAGATGCGCGGAAGAAGAG 2808
Qy      238 SerSerValThrGluPheLysSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db      2809 AGCAGTGTACAGAGTTTCATCTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTACGGAG 2868
Qy      258 IleMetSerArgProGluArgLeuLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db      2869 ATCATGTCCCGACCAAGAGGTGTTTTCATCATATTGACGGTTTCGATGACCTGGGCTCT 2928
Qy      278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProPheThr 297
Db      2929 GTCTCAACAATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAGAGAGCTCGTTACC 2988
Qy      298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db      2989 CTCATACGGCAGTCTGCTGAGGAAGTCTCTGCTCCCTGAGTCTCTTCTGATCGTCACCGTC 3048
Qy      318 ArgAspValGlyThrGluLysLeuLysSerGluValValValSerProArgTyrLeuLeuVal 337
Db      3049 AGAGAGCTGGGCACAGAGAGCTCAAGTCAGAGTCTGTCTCTCCCGTTACCTGTTAGTT 3108
Qy      338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db      3109 AGAGGAATCTCCGGGAAACAAAGAAATCCACTTGTCTCTTGAGCGCGGAGTTGGTGAGCAT 3168
Qy      358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db      3169 CAGAAGACACAAGGGTTGGTGGCATCATGAACAACCCGTGAGCTGCTCGACCAAGTGCAG 3228
Qy      378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db      3229 GTGCCCGCGTGGGCTCTCTCATCTCGTGGCCCTCGAGCTGCAGGACGTGGTGGGGGAG 3288
Qy      398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db      3289 AGCGTCGCCCTTCAACCAACCGCTCACAGGCTGCACGCCGCTTTTGTGTTTCATCAG 3348
Qy      418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db      3349 CTCACCCCTCGAGGCGTGGTGGCGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCCTGAAG 3408
Qy      438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db      3409 CGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGGAATAGGAAGTCAGTGTGTTCAGCGGTAC 3468
Qy      458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnIle 477
Db      3469 GACCTCATGTTTCAAGGACTCGGGGAGTCTGAGCTCGGTGCTCTGTTTCATATGAACATC 3528
Qy      478 LeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db      3529 CTTTCTCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTTCAGTCTCCAGGAC 3588

```

```
Qy 498 PheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 3589 TTCTGTGCGCGCTTGTAAGTCTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCGCT 3648
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 3649 CTGTACGTTGAGAAGACAAAGAGTCCATGAGGCTTAAACAGGCGAGCTTCCATATCCAC 3708
Qy 538 SerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 3709 TCGCTTTGGATGAAGCGTTTCTTGTTGGCGCTCGTAGCGGAAGACGTAAGAGGCGCACTG 3768
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTyrVal 577
Db 3769 GAGGTCTCTGCTGGGCTGTCCCGTTCCCTGGGGTGAAGCAGAAGCTTCTGCACCTGGGTC 3828
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 3829 TCTCTGTTGGGTACAGCAGCTTAATGCCACCACCCAGGAGACACCTGGAGCGCTTCCAC 3888
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 3889 TGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGCTTCCAAAGAA 3948
Qy 618 ValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 3949 GTGTGGCTTCCGATTAAACGAGAACCTGGACTTGATAGCATCTTCTTCTGCCTCCAGCAC 4008
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 4009 TGTCCGTAATTTTCGGGAAATTCGGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGTCC 4068
Qy 658 AlaGluAlaCysProValProLeuTyrMetArgAspLysThrLeuIleGluGln 677
Db 4069 GCTGAGGCAATGCTGTGTGTCCTCTATGAGTGGGGATGAAGCCCTCATTTGAGGAGCAG 4128
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 4129 TGGGAAGATTCTGCTCCATGCTTGGCACCCACCCACCTGCGGAGCTGGACCTGGGC 4188
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 4189 AGCAGCATCTGTACAGAGCGGGCCATGAAGACCCCTGTGTGCCAAGCTGAGGCATCCACC 4248
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 4249 TGCAGATACAGACCCCTGATGTTTAGAAAATGCACAGATTACCCCTGTGTGTCAGCACCTC 4308
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 4309 TGGAGAATCGTCATGGCCNACCGTAACCTAAGATCCCTCAACTTGGGAGGACCCACCTG 4368
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 4369 AAGGAAGAGGATGTAAGGATGGCGTGTGAAGCCTTAAACACCCCAAAATGTTTGTGGAG 4428
Qy 778 SerLeuArgLeuAspCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 4429 TCTTTGAGCGTGGATGCTGTGGGATGACCCATGCCCTGTACCTGGAAGATCTCCCAATC 4488
Qy 798 LeuThrThrSerProSerLeuLysSerLeuAlaGlyAsnLysValThrAspGln 817
Db 4489 CTTACGACCTCCCCAGCGCTGAATCTCTGAGCTGGCAGGAACAAAGGTGACAGCCAG 4548
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 4549 GGAGTAATGCTCTCAGTGATGCGCTTGAGAGTCTCCAGGTGCGGCTGACAGAGCTGATA 4608
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 4609 CTGAGGACTGTGGCATCACAGGCACCGGTTGCCAGAGTCTGGGCTCAGCCCTCGTCAGC 4668
```

```
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 4669 AACCGGAGCTTGACACACCTGTGCTCTATCCAAACAACAGCCCTGGGAACGAAGGTGTAAT 4728
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 4729 CTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 4788
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTyr 917
Db 4789 TGCACCTGACACGCTGGCTGTGTGTTTCTTGCACCTTTCGGCTTATGGGTAACATCATTG 4848
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 4849 CTGACGACCTGAGCCTTAGCATGAACCTGTGGAAGACAATGCGGTGAAGCTTCTGTGC 4908
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 4909 GAGGTCAATGAGAACCATCTTGTCTATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 4968
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 4969 ACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACCTGAAGAGC 5028
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 5029 CTGGATCTCAGCGACAATGCCCTGGGTGACGGTGGGGTGTGCTGACCTGTGCGAGGACTG 5088
Qy 998 LysGlnLysSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 5089 AAGCAAAAGAACAGTCTTCTGACGAGACTCGGTTTGAAGGCATGTGAGCTGACTTCTGAT 5148
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 5149 TGCTGTGAGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCACTCTGAAACCTG 5208
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 5209 GTGCAGAAATAACTTCACTCCCAAGGAATGATGAAGCTGTGTTGCGGCTTGTGCTGTCCC 5268
Qy 1058 ThrSerAsnLeuGlnIleGly 1065
Db 5269 ACCTCTAACTTACAGATAAATGGC 5292
RESULT 7
AAL47131
ID AAL47131 standard; DNA; 6939 BP.
AC AAL47131;
XX XX
DT 20-AUG-2002 (first entry)
XX XX
DE Pysin domain containing protein NALP5/Py8-hs coding sequence.
XX XX
KW Pysin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antiposrotic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; aethma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX XX
OS Unidentified.
XX XX
PN WO200240668-A2.
XX XX
PD 23-MAY-2002.
XX XX
PF 30-OCT-2001; 2001WO-EP012545.
XX XX
PR 15-NOV-2000; 2000DE-01056687.
PR 30-NOV-2000; 2000DE-01059595.
XX XX
```

PA (APOT-) APOTECH RES & DEV LTD.
XX
PI Tschopp J, Martinon F;
XX WPI; 2002-427093/45.
DR P-PSDB; AAO17859.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.
XX
XX Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x AAL47131 (1-6939)

Qy 98 IleSerGlnAlaMetGluGlnGluGluAlaAlaGluThrGluGluGlnGluIle 117
Db 3469 ATTTCACAGCTATGGAAACACAGAGGTGCCACAGCAGCAGACAGAGAAACAGAAATT 3528
Qy 118 SerGlnAlaMetGluGlnGluGluAlaAlaGluThrGluGluGlnGluHisGly 137
Db 3529 TCACAGCTATGGAAACAGAGGTGCCACAGCAGCAGACAGAGAAACAGAGCATGGA 3588
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 3589 GGTGACACATGGGACTACAGAGTCACTGATGACCAAAATTCGCTGAGGAGGAGATGA 3648
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyVala 177
Db 3649 CGTCGTAGTTTGAACACATGCTGCTGACTGGCCGGAATGCNAACGTTGGCTGGTCT 3708
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db 3709 TTTGATTCAGACCGGTGGGGCTTCGGGCTCGCACCGGTGGTCTTCGCACGGAAGTCAGGA 3768
Qy 198 IleGlyLysSerAlaLeuAlaAArgIleValLeuGluCysTrpAlaGlnGlyGlyLeuTyr 217
Db 3769 ATTGGGAAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGCGCGCAAGGTGGACTCTAC 3828
Qy 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db 3829 CAGGGAATGTTCTCTACGCTTCTCTCTCCCGTTAGAGATGTCAGCGGAGAGGAG 3888
Qy 238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db 3889 AGCAGTGTCAAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCGGCTCGGTCAGCGGAG 3948
Qy 258 IleMetSerArgProGluArgLeuPheIleAspGlyPheAspLeuGlySer 277
Db 3949 ATCATGTCCCGACAGAAAGCTGTTGTCATCATTCACGCTTTCGATGACCTGGGCTCT 4008
Qy 278 ValLeuAsnAsnAspThrLysLeuGlyAspTrpAlaGluLysGlnProProPheThr 297

Db 4009 GTCTCTCAACATGACACAAAGCTCTCAAAAGACTGGGCTGAGAGAGCAGCTCCGTTCCACC 4068
Qy 298 LeuIleArgSerIleuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db 4069 CTCTATACGCAGCTCTGCTGAGGAAGGTCTCTGCTCCCTGAGTCTCTCTGATCGTCCAGCTC 4128
Qy 318 ArgAspValGlyThrGluLysLysSerGluValValSerProArgTyrIleuLeuVal 337
Db 4129 AGAGAGCTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTCTCTCCCGCTTACCTGTAGTT 4188
Qy 338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db 4189 AGAGGAATCTCCGGGAAACAAAGATCCACTTGTCTTGAGCGCGGATTTGTTGAGCAT 4248
Qy 358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db 4249 CAGAAGACACAAAGGTTCGCTGCGATCATGAACACCGTGAGTCTCGACCAAGTCCAG 4308
Qy 378 ValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db 4309 GTGCCCGCGGTGGCTCTCTCATCTCGTGGGCCCTGCAGCTGCAGGACGTGGTGGGGAG 4368
Qy 398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db 4369 AGCGTCGCCCTTCAACCAACGCTCACAGGCTGCACGCCCTTTTGTGTTTCATCAG 4428
Qy 418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuLys 437
Db 4429 CTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCTCGAAG 4488
Qy 438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457
Db 4489 CGCTTCTGCCGTATGGCTGTGAGGAGGTGTGAATAGGAAGTCAGTGTTCAGCGTGAC 4548
Qy 458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaAlaLeuPheHisMetAsnIle 477
Db 4549 GACCTCATGTTCAAGACTTCGGGAGTCTGAGCTCCGTCGCTCTGTTTCACATGAACATC 4608
Qy 478 LeuLeuProAspSerHisCysGluGluTyrThrPhePheHisLeuSerLeuGlnAsp 497
Db 4609 CTTCTCCACAGACGACCTGTGAGGAGTACTACACCTTCTCCACCTCAGTCTCCAGGAC 4668
Qy 498 PheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db 4669 TTTCTGCGCGCTTGTACTACGTTTAGAGGGCTTGAATTCGAGCAGCTCTCTGCCCT 4728
Qy 518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db 4729 CTGTACGTTTGAGAGACAAAGAGTCCATGGAGCTTAAACAGGACGCTTCCATATCCAC 4788
Qy 538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProLeu 557
Db 4789 TCGCTTTGGATGAAGGTTTCTTGTGTGGCTCTGAGCGAAGACGTAAGAGGCGCCTG 4848
Qy 558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisIleTrpVal 577
Db 4849 GAGGTCTCTGGCTGTCTCCGTTCCCTGGGGGTGAACAGAGCTTCTGACTCGGCTC 4908
Qy 578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db 4909 TCTCTGTGGTTCAGCAGCTTAATGCCACACCCAGGAGACACCTCGAGCGCTTCCAC 4968
Qy 598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db 4969 TGTCTTTTCGAGACTCAAGACAAAGATTGTTCGCTTCGCATTAACAGCTTCCAGAA 5028
Qy 618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db 5029 GTGTGGCTTCGATTAACAGAACCTGGACTTGATAGCATCTCTCTCTCGCTCCAGCAC 5088
Qy 638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db 5089 TGTCCGTAATTTGCGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGATGCTC 5148

```
Qy 658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuLeuGluGln 677
Db 5149 GCTGAGGCACTGCTGTGCTCCCTCTATGATGCGGATAAGACCCCTCATTTGAGGAGCAG 5208
Qy 678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db 5209 TGGGAAGATTTCTGCTCATGCTTGGCACCCACCCACACCTTGGCGGAGCTGGACCTGGGC 5268
Qy 698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db 5269 AGCAGCATCTGACNAGCGGGCCATGAAGACCTTGTTGTGCAAGCTGAGGACATCCACC 5328
Qy 718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db 5329 TGCAGATACAGACCTGATGTTTAGAAATGCACAGATTACCCCTGGTGTGTCAGCACCTC 5388
Qy 738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db 5389 TGGAGAATCGTCATGGCCCAACCGTAACCTTAAGATCCCTCAACTTGGGAGGACCCACCTG 5448
Qy 758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db 5449 AAGNAGAGGATGTAAGATGCGGTGTAACCTTTAAACACCCAAATGTTTGTGGAG 5508
Qy 778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db 5509 TCTTTGAGGCTGGATTGCTGTGGATTGACCCATGCTGTTTACCTGAAGATCTCCCAATC 5568
Qy 798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db 5569 CTTTACGACCTCCCGAGCTGAAATCTCTGAGCCTGGCAGGAAACAAGGTGACAGCCAG 5628
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCTCTCAGTGATGCTTGGAGAGTCTCCACAGTCGCCCTGCAGAACTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGGAGGACTGTGGCATCACAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGGAGCTTGACACACTGTGCTATCCAAACACAGCCTGGGGAAACGAAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTGATCATGAGGCTTCCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTGGACACGGCTGGCTGTGTTCTTTCACCTTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGACCTTGAGCCTTAGCATGAACCTCTGGAAGACAATGCGTGAAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGGTGATGAGAAACCAATCTTGTCATCTCAGAGACCTTGGAGTTGGTAAAGTGTCACTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCGGCTGTGTGAGAGTCTGTCTCTGTGTATCTCGAGGACGACACCTTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGGATCTCACGGAATATGCCCTGGGTGACGCTGGGTGCTGCACATGTGCGAGGAGCTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAGAAGACAGTGTCTCAGGAGACTCGGGTTGAAGGCATGTGAGTACTGACTTCTGAT 6228
```

```
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGCACTCTCTTGGCCCTTCTGCTCAACCGGCATCTGACCACTTAAACCTG 6288
Qy 1038 ValClnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGAGAAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGTGCTGTGCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTAACTTACAGATAATTGSC 6372

RESULT 8
ID AAL47140
AAL47140 standard; DNA; 6939 BP.
XX
AC AAL47140;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP13/Py17 coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN WO200240668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01056687.
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTECH RES & DEV LTD.
XX
XX Tschopp J, Martinon F;
XX
DR WPI; 2002-427093/45.
DR P-PSDB; AAO17868.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one pyd (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 6939
Score: 968.00 Matches: 968
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.02% Indels: 0
DB: 6 Gaps: 0
```

US-10-066-521-6 (1-1344) x AAL47140 (1-6939)	
Qy	98 IleSerGlnAlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGluGlnGluIle 117
Db	3469 ATTTTCAACAAGCTATGGAACAAGAAGGTGCGACACAGCAGACAGACAGACAAGAACAAGAAATT 3528
Qy	118 SerGlnAlaMetGluGlnGluGlyAlaAlaThrAlaAlaGluThrGluGluGlnGlyHisGly 137
Db	3529 TCACAAGCTATGGAACAAGAAGGTGCGACACAGCAGACAGACAAGAACAAGGACATGGA 3588
Qy	138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluGluAspVal 157
Db	3589 GGTGACACATGGGACTACAGAGTCACTGATGATCCAAATTCGCTGAGGAGGAGATGTA 3648
Qy	158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db	3649 CGTCGTAGTTTTGAACAACACTGCTGCTGACTGGCCGCGAATGCAACGTTGGCTGGTGCT 3708
Qy	178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
Db	3709 TTTGATTCAGACCGGTGGGCTTCGGCCCTGCGACCGGTGTTCTGCAACGGAAGTCAGGA 3768
Qy	198 IleGlyLysSerAlaLeuAlaAatArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db	3769 ATTGGGAATCGGCTCTAGCCAGGAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 3828
Qy	218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
Db	3829 CAGGGAATGTTCTCTACGCTTCTCTCCCGTTAGAGAGATGCACGCGGAAGAAGAG 3888
Qy	238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db	3889 AGCAGTGTCAAGAGTTTCATCTCAGGAGTGGCGACACTCCCAAGGCTCCGGTGACGCG 3948
Qy	258 IleMetSerArgProGluArgLeuLeuPheIleAspGlyPheAspAspLeuGlySer 277
Db	3949 ATCATGTCCCGACCAAGAGCTGTGTTTCATCATTCACGCTTTCGATGACCTGGGGCTC 4008
Qy	278 ValLeuAsnAsnAspThrLysLysCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db	4009 GTCCCTCAACAATGACACAAAGCTCTGCAAGACTGGGCTGAGAAGCAGCCTCCGTTCA 4068
Qy	298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db	4069 CTCATACGCAGCTCGCTGAGGAAGTCTGCTCCCTGAGTCTCTTCCTGATCGTCACCGTC 4128
Qy	318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db	4129 AGACACGTGGGCACAGAGAAGCTCAAGTCAGAGTCTGTGCTCCCGTTTACCTGTTAGTT 4188
Qy	338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db	4189 AGAGGAATCTCCGGGAACAAGAATCCACTTGCTCTTGAGCGCGGATTTGGTGACAT 4248
Qy	358 GlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGluLeuLeuAspGlnCysGln 377
Db	4249 CAGAAGACACAAGGTTGCGTGCATCATGAACAACCGTGAGCTGCTCGACCAAGTGCAC 4308
Qy	378 ValProIaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyGlu 397
Db	4309 GTGCCCCCGTGGGCTCTCTCATCTGGTGGCCCTGCAGCTGCAAGGACGTGGGGGAG 4368
Qy	398 SerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheValPheHisGln 417
Db	4369 AGCGTCGCCCCCTTCAACCAACGCTCACAGGCTGCAACGCGCTTTTGTGTTTCATCAG 4428
Qy	418 LeuThrProArgGlyValValArgArgCysLeuAsnLeuGluArgValValLeuLys 437
Db	4429 CTCACCCCTCGAGCGTGGTCCGCGCTGTCTCAATCTGGAGGAAGAAGTTGCTCTGAAG 4488
Qy	438 ArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyAsp 457

Db	4489 CGCTTCTCCGCTATGCTGTGGAGGAGTGTGGAATAGGAAGTCAGTGTGTTGACGGTGAC 4548
Qy	458 AspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaAlaLeuPheHisMetAsnIle 477
Db	4549 GACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTTCACATGAACATC 4608
Qy	478 LeuLeuProAspSerHisCysGluGluTyrTyrPhePheHisLeuSerLeuGlnAsp 497
Db	4609 CTTTCTCCACAGACAGCAGCTGTGAGGAGTACTACACCTTCTTCCACCTCACTGTCTCCAGGAC 4668
Qy	498 PheCysAlaAlaLeuTyrTrpValLeuGluGlyLeuGluIleGluProAlaLeuCysPro 517
Db	4669 TTTCTGTCGCCCTTGATCTACGTTTAGAGGGCTCGAATTCGAGGCAGCTCTCTGCCCT 4728
Qy	518 LeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleHis 537
Db	4729 CTGTAGCTTTGAGAAGACAAAGAGTCCATGGAGCTTTAAACAGCAGGCTTCCATATCCAC 4788
Qy	538 SerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgProLeu 557
Db	4789 TCGCTTTTGGATGAAGCGTTTCTTTTGGCTCGTGAGCGAAGACGTAAGAGGCGGCATG 4848
Qy	558 GluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuHisTspVal 577
Db	4849 GAGGTCTGCTGGGCTGTCCCGTTCCTTGGGGGTGAGCAGAGGCTTCTGCACTGGGTC 4908
Qy	578 SerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheHis 597
Db	4909 TCTCTGTGGTGCAGCAGCTTAATGCCACACCCAGGAGACACCTTGGACGCTTCCAC 4968
Qy	598 CysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnGlu 617
Db	4969 TGTCTTTTCGAGACTCAAGACAAAGATTTGTTTCGCTTGGCATTAACAGCTTCCAAGAA 5028
Qy	618 ValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnHis 637
Db	5029 GTGTGGCTTCGATTAACCAAGAACCTGGACTTGATAGCATCTTCTTCTGCTTCCAGCAC 5088
Qy	638 CysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluSer 657
Db	5089 TGTCCGTATTTCGGGAAATTCGGGTGGATGTCAAAGGATCTTCCCAAGAGATGAGTCC 5148
Qy	658 AlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGln 677
Db	5149 GCTGAGGCATGCTCTGTGTCCTCTATGATGCGGGATAAGACCTTCATTGAGGAGCAG 5208
Qy	678 TrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuGly 697
Db	5209 TGGGAAGATTTCTGCTCCATGCTTGGCACCCACACCTGGCGCAGCTGGACCTGGGC 5268
Qy	698 SerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProThr 717
Db	5269 AGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCATCCACCC 5328
Qy	718 CysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisLeu 737
Db	5329 TGCAAGATCAGACCTGATGTTTAGAATGTGACAGATTACCCCTGGTGTGACGACCTC 5388
Qy	738 TrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisLeu 757
Db	5389 TGGAGAAATCGTCATGGCCAAACCGTAACCTAAGATCCCTCACTTGGGAGGCACCCACCTG 5448
Qy	758 LysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuGlu 777
Db	5449 AAGGAAGAGGATGAAGGATGGCGTGTGAAGCCTTAAACACCCAAAATGTTTGTGGAG 5508
Qy	778 SerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnIle 797
Db	5509 TCTTTGAGGCTGATGCTGTGTGGATTGACCCATGCTTACCTGAAGATCTCCCCAAATC 5568
Qy	798 LeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspGln 817
Db	5569 CTTACGACCTCCCGCAGCCTGAAATCTCTGAGCCTGCGAGGAACAAGGTGACAGACCCAG 5628

```
Qy 818 GlyValMetProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeuIle 837
Db 5629 GGAGTAATGCCTCTCAGTGATGCTTGGAGTCTCCAGTCGCGCCCTGCAGAACCTGATA 5688
Qy 838 LeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuValSer 857
Db 5689 CTGAGGAGCTGTGGCATCACAGCCACGGGTTGCCAGAGTCTGGGCTCAGCCCTCGTCAGC 5748
Qy 858 AsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyValAsn 877
Db 5749 AACCGAGCTTGACACACCTGTGCTATCCAAACAGCCCTGGGGAACGAGGTGTAAT 5808
Qy 878 LeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsnGln 897
Db 5809 CTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGATGCTGAATCAG 5868
Qy 898 CysHisLeuAspThrAlaGlyCysGlyPheLeuAlaLeuAlaLeuMetGlyAsnSerTrp 917
Db 5869 TGCCACCTGGACACGGCTGGCTGTGGTCTTCTTCACCTGGCTTATGGGTAACTCATGG 5928
Qy 918 LeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLeuCys 937
Db 5929 CTGACGACCTGAGCTTAGCATGACCTGTGGAAGCAATGCGTGAGCTTCTGTGC 5988
Qy 938 GluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHisLeu 957
Db 5989 GAGGTATGAGAGAACCATCTTGTCTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCTATCTC 6048
Qy 958 ThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLysSer 977
Db 6049 ACCGCCGCTGCTGTGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGAGC 6108
Qy 978 LeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyValAlaAlaLeuCysGluGlyLeu 997
Db 6109 CTGATCTCAGCGAATGCTTGGTACCGTGGGTGACGTGGGTGCTGCACCTGTGGAGGACTG 6168
Qy 998 LysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSerAsp 1017
Db 6169 AAGCAAAAGAACAGTGTCTCAGCAGACTCGGGTTGAAGGCATGTGCACCTCTCTGAT 6228
Qy 1018 CysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAsnLeu 1037
Db 6229 TGCTGTGAGGCACTCTCTCGGCCCTTCTCTGCAACCGGCATCTGCACGCTAAACCTG 6288
Qy 1038 ValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCysPro 1057
Db 6289 GTGCAGAAATACTTCAGTCCCAAGGAATGATGAGCTGTGTCGGCTTTGCTGTCTGCC 6348
Qy 1058 ThrSerAsnLeuGlnIleIleGly 1065
Db 6349 ACGTCTAACTTACAGATAATTGGC 6372

RESULT 9
ID ABK48628 standard; cDNA; 3900 BP.
AC ABK48628;
XX
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER cDNA.
XX
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POP;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..3603
/*tag= a
```

```
FT XX /product= "Human MATER"
PN XX WO200232955-A1.
XX XX
PD 25-APR-2002.
XX XX
PF 04-APR-2001; 2001WO-US010981.
XX XX
PR 18-OCT-2000; 2000US-0241510P.
XX XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Nelson LM, Tong Z;
DR WPI; 2002-454595/48.
DR P-PSDB; AAU79526.
XX XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX XX
PS Claim 11; Page 80-85; 93pp; English.
XX XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA.
CC Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
CC published human genomic DNA sequences, while the remainder of the
CC sequence was determined by direct cloning of human ovarian cDNAs
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: Gaps: 0

US-10-066-521-6 (1-1344) x ABK48628 (1-3900)
Qy 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
Db 568 ATTTCAAGCTATGGAAACAAAGAGGTGCCACAGCAGCAGAGACAGAACAGAAATT 627
Qy 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
Db 628 TCACAGCTATGGAAACAAAGAGGTGCCACAGCAGCAGACAGACAGACAGACATGGA 687
Qy 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
Db 688 GGTGACACATGGGACTTACAAAGCTCAGTGCATGCCAAATTCGCTGAGGAGGAGTGA 747
Qy 158 ArgArgSerPheGluAsnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
Db 748 CGTCGTAGTATTTGAAACCACTGCTGCTGAGTGGCCGGAATGCAACGTTGGCTGGTGC 807
Qy 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
```


|||||
808 TTTGATTCAGACCGTGGGGCTTCGGCCCTCGACGGTGGTTCTGCACGGAAAGTCAGGA 867
Qy
198 IleGlyLysSerAlaLeuAlaArgArgIleValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
Db
868 ATTGGGAAATCGGCTAGCCAGGAAGATCGTGTGCTGGGCGCAAGGTGGACTCTAC 927
Qy
218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysGlu 237
Db
928 CAGGGAATGTTCTCTACGCTTCTCTCCCGTTAGAGAGATGSCAGCGGAAGAAGGAG 987
Qy
238 SerSerValThrGluPheIleSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
Db
988 AGCAGTGTACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCAAGGCTCCGGTACCAGG 1047
Qy
258 IleMetSerArgProGluArgLeuPheIleIleAspGlyPheAspAspLeuGlySer 277
Db
1048 ATCATGTCCCGACCAAGAGGCTGTGTTCATCATTCAGCGTTTCGATGACCTGGGCTCT 1107
Qy
278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
Db
1108 GTCTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGGAGTGGAGAGCCCTCCGTTCA 1167
Qy
298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
Db
1168 CTCATACGAGTCTGTGAGGAAGGTCTGTCTCCCTGAGTCCCTTCCTGATCGTCACCGTC 1227
Qy
318 ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal 337
Db
1228 AGACAGTGGGCACAGAGAGCTCAAGTTCAGAGGTGTGTCTCCCGCTTACCTGTAGATT 1287
Qy
338 ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuLeuGluArgGlyIleGlyGluHis 357
Db
1288 AGAGGAATCTCCGGGGAACAAAGAAATCCACTGTCTCTTGGAGCGGGATGGTGAGCAT 1347
Qy
358 GlnLysThrGlnGlyLeuArgAlaIleMet - AsnAsnArgGluLeuLeuAspGlnCysG 377
Db
1348 CAGAAAGACACAGGGTGGCTGCATCAT - CAACACCGTGAGTGTCTGACACAGTCCCA 1406
Qy
377 nValProAlaValGlySerLeuIleCysValAlaLeuGlnLeuGlnAspValValGlyG 397
Db
1407 GGTGCCCGCGTGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGTGGTGGGGA 1466
Qy
397 uSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal - PheHisG 417
Db
1467 GAGCGTGCGCCCTTCAACCAACCGCTCACAGGCTGCAGCGCTTTTGC - GTTTCATC 1525
Qy
417 InLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL 437
Db
1526 AGCTCACCCCTCGAGGGGTGTCCGGCGCTGTCTCAATCTGGAGGAAGAAGTGTCTCTGA 1585
Qy
437 YsArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA 457
Db
1586 AGCGCTTCTGCGGTATGGCTGTGAGGAGGTGGTAATAGGAAGTCAAGTGTGTGATGGTG 1645
Qy
457 spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI 477
Db
1646 ACGACCTCATGTGTTCAAGGACTCGGGAGTCTGAGTCCGCTGTCTTTTCACATGAACA 1705
Qy
477 leLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePheHisLeuSerLeuGlnA 497
Db
1706 TCCTTCTCCACAGACGACCTGTGAGGAGTACTACACTTCTTCCACTCAGTCTCAGG 1765
Qy
497 spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP 517
Db
1766 ACTTCTGTGCGCCCTGTACTACGTGTTAGAGGCTTGAAATCGAGCCAGCTCTCTGCC 1825
Qy
517 roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH 537
Db
1826 CTCGTGACGTGGAAGACAAAGAGGTCCATGGAGCTTAAACAGGAGGCTTCCATATCC 1885
Qy
537 isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL 557
|||||

Db
1886 ACTCGCTTTGGATGAAGCGTTTCTTGTGGCTCGTAGCGAAGACGTAAAGAGAGCCAC 1945
Qy
557 euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpY 577
Db
1946 TGGAGTCTCTGTGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAAGCTTCTGCACCTGG 2005
Qy
577 alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH 597
Db
2006 TCTCTCTGTGGGTACAGACCTTAATGCCACCAACCCAGGAGACACCTGGAGCCCTTCC 2065
Qy
597 isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG 617
Db
2066 ACTGTCTTTTCAGACTCAAGACAAAGAGTTGTTCGCTTGGCATTAACAGAGTTCAAG 2125
Qy
617 luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH 637
Db
2126 AAGTGTGGCTTCGATTAACACAGAACCTGGACTTGATAGCATCTTCTCTCTGCCCTCCAG 2185
Qy
637 isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS 657
Db
2186 ACTGTCCGTATTTTCGGGAAAATTCGGGTGGATGTCAAAGGGATCTTTCCCAAGAGATGAGT 2245
Qy
657 exAlaGluAlaCysProValValProLeuTrpMetArgAspLysThrLeuIleGluGluG 677
Db
2246 CCGCTGAGGATGTCTGTGTCTCTATGATGCGGATAGACCCCTCATTTGAGAGGC 2305
Qy
677 InTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG 697
Db
2306 AGTGGGAAGATTCTCTCATGTCTGGCACCCACCCACACCTGCGGACGTGGACCTGG 2365
Qy
697 lysSerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProT 717
Db
2366 GCAGCAGCATCTTCAGACGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCA 2425
Qy
717 hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL 737
Db
2426 CCTGCAAGATACAGACCTGTATGTTTGAATAATGCACAGATTACCCCTGGTGTGCACACC 2485
Qy
737 euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL 757
Db
2486 TCTGGAGAATCGTCATGGCCAAACCGTAACCTAAGATCCCTCAACTTGGGAGGCACCCACC 2545
Qy
757 euLysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG 777
Db
2546 TGAAGGAAGAGGATGAAGATGGCGTGTGAAGCCCTTAAACACACCCAAAATGTTTGTGG 2605
Qy
777 luSerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI 797
Db
2606 AGTCTTTGAGGCTGGATTGCTGTGGATTGCCCATGCCCTGTACCTGAAGATCTCCAAA 2665
Qy
797 leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG 817
Db
2666 TCCTTACGACCTCCCCAGCTGAAATCTCTGAGCCCTGCGAGGAAACAAGGTGACAGACC 2725
Qy
817 InGlyValMet - ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuLysLeu 836
Db
2726 AGGAGTAAC - GCCTCTCAGTATGCTTGAAGGCTCTCCAGTGCCTCTGCAAGAAGCTG 2784
Qy
837 IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal 856
Db
2785 ATACTGGAGGACTGTGGCATCACAGCCAGGGGTGGCAGAGTCTGGCCTCAGCCCTCGTC 2844
Qy
857 SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGluGlyVal 876
Db
2845 AGCAACCGGAGCTTGACACACCTGTGCCTATCAACAACAGCCTGGGGAAACGAAGGTGA 2904
Qy
877 AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn 896
Db
2905 AATCTACTGTGTGCATCCATGAGGCTTCCCACTGTAGTCTGCAGAGGCTGTATGCTGAAT 2964
Qy
897 GlnCysHisLeuAspThrAlaGlyCysGlyPhe - LeuAlaLeuAlaLeuMetGlyAsnSe 916
Db
2965 CAGTGCACCTGAGACGCGCTGGCTGTGG - TTCTCTTGCACTTGCGCTTATGGGTAATC 3023


```
QY 916 rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyVallysLeuLe 936
DB 3024 ATGGCTGACGACCTTAGCCTTAGCATCAACCTGTGAGACCAATGGCGTGAAGCTTCT 3083
QY 936 uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuVallysCysHi 956
DB 3084 GTGCCAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCTGGAGTTGGTAAAGTGTCA 3143
QY 956 sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy 976
DB 3144 TCTACCCGCGGTGCTGTGAGAGTCTGTCTGTGTATCTGAGGACACACCTGAA 3203
QY 976 sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluGl 996
DB 3204 GAGCCTGGATCTCACGGACAATGCCCTGGGTGACGGTGGGTGTGCTGGCTGTGGAGGG 3263
QY 996 yLeuLysGlnLysAnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe 1016
DB 3264 ACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCATGTGGACTGACTTC 3323
QY 1016 rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs 1036
DB 3324 TGATGTCTGTGAGGACATCTCTTGGCCCTTCTTGCACCCGGCATCTGACCACTCTAAA 3383
QY 1036 nLeuValGlnAnAnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy 1056
DB 3384 CCTGGTGCAATAACTTCACTCCCAAGGAATGATGAAGCTGTGTTCGGCCTTTGCTG 3443
QY 1056 sProThrSerAnLeuGlnIleGlyLeuTrpLysTrpGlnTrpValGlnIleAr 1076
DB 3444 TCCCACTGTGAGGAGTACAGATAATGGGCTGTGGAAATGGCAGTACCCCTGTGCAATAAG 3503
QY 1076 gLysLeuLeuGluGluValGlnLeuLeuLysProArgValValIleAspGlySerTrpHi 1096
DB 3504 GAAGCTGTGGAGGAGTGCAGCTACTCAGCCCCGAGTGTGAATGACGTAATGACGTAATGGCA 3563
QY 1096 sSerPheAspGluAspArg 1103
DB 3564 TTCCTTTGATGAAGATGACCGG 3585
RESULT 10
AAD49018
ID AAD49018 standard; cDNA; 3900 BP.
XX AC AAD49018;
XX DT 07-MAR-2003 (first entry)
XX DE Human MATER cDNA.
XX KW Human; MATER protein; infertility; fertility; contraceptive agent;
XX KW gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..3603
XX FT /*tag= a
XX FT /product= "Human MATER protein"
XX PN WO200281492-A1.
XX PD 17-OCT-2002.
XX PF 29-MAR-2002; 2002WO-US009776.
XX PR 04-APR-2001; 2001WO-US010981.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nelson LM, Tong Z;
XX XX
```

```
DR WPI; 2003-058494/05.
XX P-PSDB; AAE31749.
PT New isolated variant MATER proteins and nucleic acids, useful for
PT diagnosing, prognosing and treating infertility and reduced fertility,
XX and as contraceptive agents.
XX Example 2; Page 93-98; 110pp; English.
XX The present invention relates to novel MATER proteins and polynucleotides
CC encoding such proteins. The MATER proteins are essential to fertility.
CC Sequences of the invention are useful for diagnosing, prognosing and
CC treating infertility, reduced fertility and as contraceptive agents. They
CC are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
CC infertility or reduced fertility. The present sequence is human MATER
CC cDNA
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3900
Score: 616.00 Matches: 1002
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 4
Query Match: 45.83% Indels: 8
DB: 7 Gaps: 0

US-10-066-521-6 (1-1344) x AAD49018 (1-3900)
QY 98 IleSerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGluIle 117
DB 568 ATTTCAACAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGAGAACAAGAAATT 627
QY 118 SerGlnAlaMetGluGlnGluGlyAlaThrAlaAlaGluThrGluGlnGlyHisGly 137
DB 628 TCAAGAGCTATGGAACAAGAAGGTGCCACAGCAGCAGACAGAGAACAAGACATGGA 687
QY 138 GlyAspThrTrpAspTyrLysSerHisValMetThrLysPheAlaGluGluAspVal 157
DB 688 GGTGACACATGGGACTACAGAGTCAAGAGTCAAGTATGACCAATTCGCTGAGGAGGAGTGA 747
QY 158 ArgArgSerPheGluAnThrAlaAlaAspTrpProGluMetGlnThrLeuAlaGlyAla 177
DB 748 CGTCGTAGTTTTCGAAACACTGTCTGCTGACTGCGCGGAATGCAACGTTGGCTGCT 807
QY 178 PheAspSerAspArgTrpGlyPheArgProArgThrValValLeuHisGlyLysSerGly 197
DB 808 TTTGATTACAGACCGTGGGGCTTCCGGCTTCGCACGCTGCTGCGCAAGTTCAGGA 867
QY 198 IleGlyLysSerAlaLeuAlaArgAlaValLeuCysTrpAlaGlnGlyGlyLeuTyr 217
DB 868 ATTTGGGAATCGGCTCTAGCCAGAAGATCGTCTGTCTGGCGGCAAGTGGACTCTAC 927
QY 218 GlnGlyMetPheSerTyrValPhePheLeuProValArgGluMetGlnArgLysLysGlu 237
DB 928 CAGGGAATGTTCTCTACCGTCTTCTTCTCCCTCCCGTTAGAGAGATGCAGCGAAGAGGAG 987
QY 238 SerSerValThrGluPheLeuSerArgGluTrpProAspSerGlnAlaProValThrGlu 257
DB 988 AGCAGTGTACAGAGTTTCATCTCCAGGAGTGGCCAGACTCCCGAGGCTCCGCTGACGAG 1047
QY 258 IleMetSerArgProGluArgLeuLeuPheIleAspGlyPheAspLeuGlySer 277
DB 1048 ATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGGTTTCGATGACCTGGGCTCT 1107
QY 278 ValLeuAsnAsnAspThrLysLeuCysLysAspTrpAlaGluLysGlnProProPheThr 297
DB 1108 GTCTCTCAACAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGCAGCTCCGTTCA 1167
QY 298 LeuIleArgSerLeuLeuArgLysValLeuLeuProGluSerPheLeuIleValThrVal 317
DB 1168 CTCATACGCGAGTCTGCTGAGGAAGGTCTGTCTCCCTGAGTCTTCTCTGATCGTCACCGTC 1227
```

QY	318	ArgAspValGlyThrGluLysLeuLysSerGluValValSerProArgTyrLeuLeuVal	337
Db	1228	AGAGCGTGGGCACAGAGAAGCTCAAGTCAGAGGTCTGTGTCTCCCGTTTACCTGTTAGTT	1287
QY	338	ArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGluArgGlyIleGlyGluHis	357
Db	1288	AGAGGAATCTCCGGGAAACAAGAATTCACCTTGCTCTTGAGCCGGGATTTGTGACAT	1347
QY	358	GlnLysThrGlnGlyLeuArgAlaIleMet-AsnAsnArgGluLeuLeuAspGlnCysG1	377
Db	1348	CAGAAGACACAAAGGTTGCGTGCATCAT-CAACAACCGTGAGCTCTCGACCACTGCCA	1406
QY	377	nValProAlaValGlySerIleCysValAlaLeuGlnLeuGlnAspValValGlyG1	397
Db	1407	GGTCCCGCGGTGGCTCTCTCATCTCGTGGCCCTGCAGCTGCAGGACGTGTGGGGGA	1466
QY	397	userValAlaProPheAsnGlnThrLeuThrGlyLeuHisAlaAlaPheVal-PheHisG	417
Db	1467	GAGCGTGGCCCTTCAACCAACGCTCACAGGCTGCAGCCGCTTTGCG- GTTTCATC	1525
QY	417	lnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGluGluArgValValLeuL	437
Db	1526	AGCTCACCCCTCGAGGCGTGTCCGCGCTGTCTCAATCTGGAGGAAGAGTTGTCTCGA	1585
QY	437	ysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLysSerValPheAspGlyA	457
Db	1586	AGCGCTTCTGCGGTATGGCTGTGAGGAGGTGGAAATAGGAAGTCAGTGTTTGATGGTG	1645
QY	457	spAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAlaLeuPheHisMetAsnI	477
Db	1646	ACGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGNACA	1705
QY	477	leLeuLeuProAspSerHisCysGluGlyTyrTyrThrPhePheHisLeuSerLeuGlnA	497
Db	1706	TCCTTCTCCACAGACGACCTGTGAGGAGTACTACACCTTCTCCACCTCAGTCTCCAGG	1765
QY	497	spPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIleGluProAlaLeuCysP	517
Db	1766	ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGGAAATCGAGCCAGCTCTCTGCC	1825
QY	517	roLeuTyrValGluLysThrLysArgSerMetGluLeuLysGlnAlaGlyPheHisIleH	537
Db	1826	CTCTGTACGTTGAGAACAAAGAGTCCATGGAGCTTAAACAGCAGGCTTCCATATCC	1885
QY	537	isSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGluAspValArgArgProL	557
Db	1886	ACTCGCTTTGGATGAAGCGTTTCTGTTTGCCCTCGTGAGCGAAGACGTAAAGAGGCCAC	1945
QY	557	euGluValLeuLeuGlyCysProValProLeuGlyValLysGlnLysLeuLeuHisTrpV	577
Db	1946	TGGAGGTCTCTGCGGTGTCCTGTTCCCTTGGGGGTGAAGACGAGCTTCTGCACTGGG	2005
QY	577	alSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAspThrLeuAspAlaPheH	597
Db	2006	TCTCTCTGTTGGGTACAGAGCTAATGCCACCACCCCGAGGAGACACCTTGAGCGCTTCC	2065
QY	597	isCysLeuPheGluThrGlnAspLysGluPheValArgLeuAlaLeuAsnSerPheGlnG	617
Db	2066	ACTGCTCTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTTAAACAGACTTCCAAG	2125
QY	617	luValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSerSerPheCysLeuGlnH	637
Db	2126	AAGTGTGGCTTCCGATTAACAGAACTGGACTGTGATAGCATCTTCTCTGCGCTCCAGC	2185
QY	637	isCysProTyrLeuArgLysIleArgValAspValLysGlyIlePheProArgAspGluS	657
Db	2186	ACTGTCGTAATTCGGGAAATTCGGGTGGATGTCAAGGGATCTTCCCAAGAGATGAGT	2245
QY	657	erAlaGluAlaCysProValProLeuTrpMetArgAspLysThrLeuIleGluG1	677
Db	2246	CCGCTGAGGATGTCTGTGTCCTCTATGATGTCGGGATGAAGACCTCAATTGAGGAGC	2305

QY	677	lnTrpGluAspPheCysSerMetLeuGlyThrHisProHisLeuArgGlnLeuAspLeuG	697
Db	2306	AGTGGGAAGATTTCCTCATGCTTGGCACCACCACACCTTGGCCGACGTGCAGCTGG	2365
QY	697	lySerSerIleLeuThrGluArgAlaMetLysThrLeuCysAlaLysLeuArgHisProF	717
Db	2366	GCAGCAGCATCTGCAGAGCGGCCCATGAAGACCTGTGTGCCAAGCTGAGGCATCCCA	2425
QY	717	hrCysLysIleGlnThrLeuMetPheArgAsnAlaGlnIleThrProGlyValGlnHisL	737
Db	2426	CCTGCAAGATACAGACCTGTATGTTTAGAAATGCACAGATTACCCCTGGTGTCAACACC	2485
QY	737	euTrpArgIleValMetAlaAsnArgAsnLeuArgSerLeuAsnLeuGlyGlyThrHisL	757
Db	2486	TCGGGAATTCGTCATGGCCAAACCCGTAACCTAAGATCCCTCAACTTGGGAGGCACCACC	2545
QY	757	eulysGluGluAspValArgMetAlaCysGluAlaLeuLysHisProLysCysLeuLeuG	777
Db	2546	TGNAAGGAGGATGTAAAGATGGCGTGTGAAGCTTTAAACACCCCAAAATGTTTGTGG	2605
QY	777	lySerLeuArgLeuAspCysCysGlyLeuThrHisAlaCysTyrLeuLysIleSerGlnI	797
Db	2606	AGTCTTTGAGGCTGGATTGCTGTGGATTGACCATGCTGTGTACCTGAAAGATCTCCCAA	2665
QY	797	leLeuThrThrSerProSerLeuLysSerLeuSerLeuAlaGlyAsnLysValThrAspG	817
Db	2666	TCCTTACGACCTCCCCAGCTGAAATCTCTGAGCCTGGCAGGAACAAGGTGACAGACC	2725
QY	817	lnGlyValMet-ProLeuSerAspAlaLeuArgValSerGlnCysAlaLeuGlnLysLeu	836
Db	2726	AGGGAGTAAC-GCCTCTCAGTGATGCTTGGAGGCTCTCCAGTGCGCCCTGCAGAACTG	2784
QY	837	IleLeuGluAspCysGlyIleThrAlaThrGlyCysGlnSerLeuAlaSerAlaLeuVal	856
Db	2785	ATACTGGAGGACTGTGGCATCACAGCACGGGTGCGAGAGTCTGGCCTCAGCCCTCGTC	2844
QY	857	SerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSerLeuGlyAsnGlyVal	876
Db	2845	AGCAACCGAGCTTGACACACCTTGCCCTATCCAAACAGCCTTGGGGAACGAAGGTGA	2904
QY	877	AsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeuGlnArgLeuMetLeuAsn	896
Db	2905	AATCTACTGTGTCGATCCATGAGGCTTCCCATGCTAGTCTGCAGAGGCTGATGCTGAAT	2964
QY	897	GlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLeuAlaLeuMetGlyAsnSe	916
Db	2965	CAGTGCACCTGGACACGGCTGCTGTGG-ITCTCTTGCACTTGGCTTTATGGGTAACTC	3023
QY	916	rTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAspAsnGlyValLysLeuLe	936
Db	3024	ATGGCTGACGCACTGAGCCTTAGCATGAACCTGTGGGAAGACAATGGCGTGAGCTTCT	3083
QY	936	uCysGluValMetArgGluProSerCysHisLeuGlnAspLeuGluLeuValLysCysHi	956
Db	3084	GTGCGAGGTCATGAGAGAACCATTCTGTCTATCTCCAGGACCTGGAGTGGTAAAGTGTCA	3143
QY	956	sLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSerArgSerArgHisLeuLy	976
Db	3144	TCTACCGCGCGTGTGTGAGAGTCTGCTGTGTGATCTCGAGGAGCAGACACCTGAA	3203
QY	976	sSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyValAlaAlaLeuCysGluG1	996
Db	3204	GAGCCTGGATCTCACGGACATGCCCTGGGTGACGTGGGTGGGTGGCTGTGCGAGGG	3263
QY	996	yLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLysAlaCysGlyLeuThrSe	1016
Db	3264	ACTGAAGCAAAAGAACAGTGTCTGTACGAGACTCGGGTTGAAGGCATGTGGACTGACTTC	3323
QY	1016	rAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnArgHisLeuThrSerLeuAs	1036
Db	3324	TGATTCGTGTGAGGCACTCTCCTTGGCCCTTTCTGTCGAACCGGCATCTCACCAGTCTAA	3383
QY	1036	nLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLeuCysSerAlaPheAlaCy	1056

Db 3384 CTGTGTGCGAATAACTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGGCCGTG 3443
Qy 1056 sProThrSerAsnLeuGlnIleilegLyLeuTrpLysTrpGlnTyrProValGlnIleAr 1076
Db 3444 TCCACGCTTAACCTACAGATTAATTTGGGCTGTGGAAATGGCAGTACCTCTGTGCAAAATAG 3503
Qy 1076 gLysLeuLeuGluValGlnLeuLysProArgValValIleAspGlySerTrpHi 1096
Db 3504 GAAAGCTGTGGAGGAAGTGCAGTACTCAAGCCCGAGTTCGTAATTGACGGTAGTTGCA 3563
Qy 1096 sSerPheAspGluAspAspArg 1103
Db 3564 TTCTTTTGTATGAAGATGACCCG 3585

RESULT 11

ABK48609
ID ABK48609 standard; cDNA; 1157 BP.
AC ABK48609;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER cDNA fragment #1.
XX
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW Maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1155
FT /*tag= a
FT /product= "Human MATER"
FT /partial
FT /note= "No start or stop codon shown"
XX
PN WO200232955-A1.
XX
XX 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
XX 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI: 2002-454595/48.
DR P-PSDB; AAU79523.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 11; Page 68-69; 93pp; English.

The invention discloses an isolated human MATER (Maternal Antigen That Embryos Require) protein, which can complement a MATER null phenotype in which zygotes arising from the oocyte do not progress beyond the two-cell stage. MATER is required for early embryonic survival and abnormal levels of the protein can lead to premature ovarian failure (POF) and can be caused by under or over expression of MATER or an autoimmune response to MATER. MATER is a single-copy Maternal effect gene found on chromosome 19. The MATER protein and polynucleotide, by gene therapy, are useful in diagnosing or treating fertility and reduced fertility. In particular, the MATER protein is useful as a contraceptive agent, or for influencing (either inhibiting or enhancing) fertility and can be used to detect a predisposition to infertility or reduced fertility, or for presymptomatic screening of an individual for infertility/reduced fertility. The protein

CC and polynucleotide are also useful for detecting an excess or deficiency, or genetic mutation, of the MATER protein in a mammalian subject (e.g. a human or a mouse) or for screening for a compound useful in influencing CC MATER-mediated fertility. The sequence presented is the human MATER cDNA CC fragment #1, which was isolated from a human cDNA library
XX

SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 1157
Score: 385.00 Matches: 385
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.65% Indels: 0
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x ABK48609 (1-1157)

Qy 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleAspGly 271
Db 1 CAAGCTCCGGTGACGGAGATCATGTCCGACCAAGAGGCTGTTGTTTCATCATTTGACGGT 60
Qy 272 PheAspLeuGlySerValLeuAsnAspThrLysLeuCysLysAspTrpAlaGlu 291
Db 61 TTCGATGACCTGGGCTCTGTCTCAACAATGACAAAGCTCTGCAAGACTGGGCTGAG 120
Qy 292 LysGlnProPheThrLeuIleArgSerLeuLeuArgLysValLeuLeuProGluSer 311
Db 121 AAGCAGCCTCCGTTACCTCATACCCCTCATACGAGTCTGTCTGAGGAAGGTCCTGCTCCCTGAGTCC 180
Qy 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331
Db 181 TTCCTGATCGTCACCGTCAGAGACGTGGGCACACAGAAAGCTCAAGTCAGAGCTGTGTCT 240
Qy 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu 351
Db 241 CCCCGTTACCTGTAGTTAGAGGAATCTCCGGGGAACAAGAAATCCACTTCTCTCTTCTGAG 300
Qy 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371
Db 301 CGCGGGATTGTGAGCATCAGAACAAAGGTTGCGTGCATCATGAACACCGTGAG 360
Qy 372 LeuLeuAspGlnCysGlnValProAlaValIcLySerLeuIleCysValAlaLeuGlnLeu 391
Db 361 CTGCTCGACAGTCCAGGTGCCCGCGTGGGCTCTCTCATCTCGGTGGCCCTCAGCTG 420
Qy 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411
Db 421 CAGGACGTGGTGGGGAGAGCGTCGCCCTTCAACCAAAACGCTCACAGGCTGCACGCC 480
Qy 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431
Db 481 GCITTTTGTGTTTCATCAGCTCACCCCTCGAGGGCGTGGTCCGGCGCTGTCTCAATCTGGAG 540
Qy 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTrpAsnArgLys 451
Db 541 GAAAGAGTTGTCTGAAGCGCTTCTGCGGTATGCGCTGTGGAGGGAGTGTGGAAATAGGAG 600
Qy 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471
Db 601 TCAGTGTTCAGCGGTGACGACCTCATGGTTCAGGACTCCGGGAGTCTGAGCTCCGTGCT 660
Qy 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePhe 491
Db 661 CTGTTTTCACATGAACATCTCTTCTCCAGACAGCCACTGTGTGAGGAGTACTACACCTTCTTC 720
Qy 492 HisLeuSerLeuGlnAspPheCysAlaAlaLeuTyrTyrValLeuGluGlyLeuGluIle 511
Db 721 CACCTCAGTCTCCAGGACTCTCTGTGCGCGCTTGTACTACGTGTAGAGGGCTCGGAATC 780
Qy 512 GluProAlaLeuCysProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAAGAAAGAGGTCCATGGAGCTTTAAACAG 840

QY 532 AlaGlyPheHisIleHisSerLeuTyrMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGCITTCATATCCACTCGCTTGGATGAAGCGTTTCCTTGGCTCGTGAAGCAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GACGTAAAGAGCCNACTGGAGGTCTGCTGGGCTGTCCCGTCCCTGGGGGTGAGCAG 960
QY 572 LysLeuLeuHisTyrValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGCACCTGGCTCTCTGTGGGTGAGCAGCCTAATGCCACCCAGGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGGAGCGCTTCCACTGCTCTTTTCAGACTCAAGACAAAGAGTTGTTCGCTTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTyrLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAAACAGCTTCCAGAGTGTGGCTCCGATTACCAAGACCTGGACTTGTATGATCT 1140
QY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155

RESULT 12

AAD48999
ID AAD48999 standard; cDNA; 1157 BP.

XX

AC AAD48999;

XX

DT 07-MAR-2003 (first entry)

XX

DE Human MATER cDNA fragment 1.

XX

KW Human; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS

1..1155
/*tag= a
/product= "Human MATER protein fragment 1"
/transl_except= (pos:112..114, aa:Thr)
/transl_except= (pos:589..591, aa:Thr)
/transl_except= (pos:865..867, aa:Thr)
/transl_except= (pos:973..975, aa:Thr)
/transl_except= (pos:1102..1104, aa:Thr)
/note= "No start and stop codon"
/partial

FT

FT

XX

PN WO200281492-A1.

XX

PD 17-OCT-2002.

XX

PF 29-MAR-2002; 2002WO-US009776.

XX

PR 04-APR-2001; 2001WO-US010981.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Nelson LM, Tong Z;

XX

XX WPI; 2003-058494/05.

DR

P-PSDB; AAE31746.

XX

XX New isolated variant MATER proteins and nucleic acids, useful for

PT

PT diagnosing, prognosing and treating infertility and reduced fertility,

PT

PT and as contraceptive agents.

XX

PS Example 2; Page 79-80; 110pp; English.

CC The present invention relates to novel MATER proteins and polynucleotides
CC encoding such proteins. The MATER proteins are essential to fertility.
CC Sequences of the invention are useful for diagnosing, prognosing and
CC treating infertility, reduced fertility and as contraceptive agents. They
CC are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
CC infertility or reduced fertility. The present sequence is human MATER
CC cDNA fragment 1

XX

SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	1157
Score:	385.00	Matches:	385
Score Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	28.65%	Indels:	0
DB:	7	Gaps:	0

US-10-066-521-6 (1-1344) x AAD48999 (1-1157)

QY 252 GlnAlaProValThrGluIleMetSerArgProGluArgLeuLeuPheIleLeuAspGly 271

Db

1 CAAGCTCCGGTGACGGAGATCATGTCCGACCAAGAGGCTGTGTTCATCATTTGACGGT 60

QY 272 PheAspAspLeuGlySerValLeuAsnAsnAspThrLysLeuCysLysAspTyrAlaGlu 291

Db

61 TTCGATGACCTGGGCTCTGTCTCAACATGACACAAAGCTCTGCAAGACTGGGCTGAG 120

QY 292 LysGlnProPheThrLeuIleArgSerLeuArgLysValLeuLeuProGluSer 311

Db

121 AAGCAGCTCCGTTCCCTCATACGAGTCTCTGTAGGAAGTCTCTGCTCCCTGAGTCC 180

QY 312 PheLeuIleValThrValArgAspValGlyThrGluLysLeuLysSerGluValValSer 331

Db

181 TTCCTGATGTCACCGTCAGAGAGCTGGGCACAGAGAGCTCAAGTCAGAGGTCGTGCT 240

QY 332 ProArgTyrLeuLeuValArgGlyIleSerGlyGluGlnArgIleHisLeuLeuGlu 351

Db

241 CCCCGTTACCTGTAGTGTAGAGGAATCTCCGGGGAACAAGAAATCCACTTGTCTCTG 300

QY 352 ArgGlyIleGlyGluHisGlnLysThrGlnGlyLeuArgAlaIleMetAsnAsnArgGlu 371

Db

301 CGCGGATTTGTCAGCATCAGAGACACAGGGTTCGTCGATCATGAACACCGTGAG 360

QY 372 LeuLeuAspGlnCysGlnValProAlaValGlySerLeuIleCysValAlaLeuGlnLeu 391

Db

361 CTGCTCGACAGTCCAGGTCGCCCGCTGGGCTCTCTCATCTCGTGGCCCTGCAGCTG 420

QY 392 GlnAspValValGlyGluSerValAlaProPheAsnGlnThrLeuThrGlyLeuHisAla 411

Db

421 CAGGACGTGGTGGGGAGAGCGTCGCCCTTCAACCAACCGCTCAGAGCTGCAGGCC 480

QY 412 AlaPheValPheHisGlnLeuThrProArgGlyValValArgArgCysLeuAsnLeuGlu 431

Db

481 GCITTTGTGTTTCATCAGCTCACCCCTCAGAGCGTGTCCGGCGCTGTCTCAATCTGGAG 540

QY 432 GluArgValValLeuLysArgPheCysArgMetAlaValGluGlyValTyrAsnArgLys 451

Db

541 GAAAGAGTTCTCTGAAGCGCTTCTCCGTATGGCTGTGGAGGGAGTGTGGAATAGGAAG 600

QY 452 SerValPheAspGlyAspAspLeuMetValGlnGlyLeuGlyGluSerGluLeuArgAla 471

Db

601 TCAGTGTTCAGCGTGACGACCTCATGGTTCAGAGACTCGGGGAGTGTGAGCTCCGTGCT 660

QY 472 LeuPheHisMetAsnIleLeuLeuProAspSerHisCysGluGluTyrTyrThrPhePhe 491

Db

661 CTGTTTCACATGAACATCTCTCCAGACAGCCACTGTGAGGAGTACTACACTCTTCTTC 720

QY 492 HisLeuSerLeuGlnAspPheCysAlaLeuTyrTyrValLeuGluGlyLeuGluIle 511

Db

721 CACCTCAGTCTCCAGGACTTCTGTGCGCGCTTGTACTACGTGTAGAGGGCGCTGAATC 780

XX

QY 512 GluProAlaLeuCyProLeuTyrValGluLysThrLysArgSerMetGluLeuLysGln 531
Db 781 GAGCCAGCTCTCTGCCCTCTGTACGTTGAGAGACAAAGAGGTCATGGAGCTTAAACAG 840
QY 532 AlaGlyPheHisIleHisSerLeuTrpMetLysArgPheLeuPheGlyLeuValSerGlu 551
Db 841 GCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTGGGCTCGTGAGCGAA 900
QY 552 AspValArgArgProLeuGluValLeuLeuGlyCysProValProLeuGlyValLysGln 571
Db 901 GAGGTAAAGAGGCCACTGGAGGTCCTGCTGGGCTCTCCGTTCCCTGGGGGTGAAGCAG 960
QY 572 LysLeuLeuHisTrpValSerLeuLeuGlyGlnGlnProAsnAlaThrThrProGlyAsp 591
Db 961 AAGCTTCTGCATGGGTCTCTGTGGGTGACGAGCCTAATGCCACCACCCAGGAGAC 1020
QY 592 ThrLeuAspAlaPheHisCysLeuPheGluThrGlnAspLysGluPheValArgLeuAla 611
Db 1021 ACCCTGGAGCGCTTCCACTGCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCA 1080
QY 612 LeuAsnSerPheGlnGluValTrpLeuProIleAsnGlnAsnLeuAspLeuIleAlaSer 631
Db 1081 TTAACAGCTTCCAGAGAGTGTGGCTTCGATTAAACGAGACCTGGACTTGTATGACTCT 1140
QY 632 SerPheCysLeuGln 636
Db 1141 TCCTTCTGCTCCAG 1155
RESULT 13
ABK48610
ID ABK48610 standard; cDNA; 1075 BP.
XX
AC ABK48610;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER cDNA fragment #2.
XX
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 2..778
FT /*tag= a
FT /product= "Human MATER"
FT /partial
FT /note= "No start codon shown"
XX
FN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
XX WPI; 2002-454595/48.
DR P-P8DB; AAU79524.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 11; Page 70-71; 93pp; English.
XX

The invention discloses an isolated human MATER (Maternal Antigen That Embryos Require) protein, which can complement a MATER null phenotype in which zygotes arising from the oocyte do not progress beyond the two-cell stage. MATER is required for early embryonic survival and abnormal levels of the protein can lead to premature ovarian failure (POF) and can be caused by under or over expression of MATER or an autoimmune response to MATER. MATER is a single-copy maternal effect gene found on chromosome 19. The MATER protein and polynucleotide, by gene therapy, are useful in diagnosing or treating fertility and reduced fertility. In particular, the MATER protein is useful as a contraceptive agent, or for influencing (either inhibiting or enhancing) fertility and can be used to detect a predisposition to infertility or reduced fertility, or for presymptomatic screening of an individual for infertility/reduced fertility. The protein and polynucleotide are also useful for detecting an excess or deficiency, or genetic mutation, of the MATER protein in a mammalian subject (e.g. a human or a mouse) or for screening for a compound useful in influencing human-mediated fertility. The sequence presented is the human MATER cDNA fragment #2, which was isolated from a human cDNA library

XX SQ Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.46e-184 Length: 1075
Score: 196.00 Matches: 252
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 6 Gaps: 0

US-10-066-521-6 (1-1344) x ABK48610 (1-1075)

QY 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnSer 870
Db 2 CTGGCCTCAGCCCTCGTCAGCAACCGGAGCTTGACACACCTGTGCTTATCCACACAGC 61
QY 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAAGGTGTAATCTACTGTGTCATCCATGAGGCTTCCCTAGTGTCTG 121
QY 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCTGGACACGGCTGGCTGTGG-TTCTCTTGGCACT 180
QY 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGCCTTTATGGGTAACTCATGGCTGACGCACCTGAGGCTTAGCATGAACCTGTGGAGA 240
QY 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTCTATGAGAGAACCATCTTGTTCATCTCCAGGACCT 300
QY 950 uGluLeuValLysCysHisLeuThrAlaAlaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTTGGTAAATGTCTATCCACCGCGGTGCTGTGAGAGTCTGTCTGTGTGATCTC 360
QY 970 rArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGAGCTTGGATCTCAGCGACAAATGCCCTGGGTGAGCGTGGGT 420
QY 990 lAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAA 480
QY 1010 sAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCAATGTGGACTGACTTCTGATTGCTGTGAGGCACTCTCTCTTGGCCCTTTCTCTCAACCG 540
QY 1030 qHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACAGTCTAAACCTGGTGCAAGATACTTCAGTCCCAAGGAATGATGAAGCT 600
QY 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTrpGl 1070

Db 601 GTGTTGGCCCTTGCTGTCCACGCTAACTCAGATAAATGGGCTGTGAAATGSCA 660
Qy 1070 nTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArgValVa 1090
Db 661 GTACCTGTGCAATPARGAGCTGCTCGAGGAAGTGACGCTACTCAAGCCCGAGTGT 720
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTTGGCAATCTTTTGATGAAGATGACCG 760

RESULT 14

AAAD49000
ID AAD49000 standard; cDNA; 1075 BP.

XX AAD49000;

XX 07-MAR-2003 (first entry)

XX Human MATER cDNA fragment 2.

XX Human; MATER protein; infertility; fertility; contraceptive agent;
KW gene therapy; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..778
FT /*tag= a
FT /product= "Human MATER protein fragment 2"
FT /transl_except= (pos:200..202, aa:Thr)
FT /transl_except= (pos:650..652, aa:Thr)
FT /transl_except= (pos:656..658, aa:Thr)
FT /transl_except= (pos:734..736, aa:Thr)
FT /transl_except= (pos:764..766, aa:Thr)
FT /transl_except= (pos:767..769, aa:Thr)
FT /note= "No start codon"
FT /partial

XX W0200281492-AL.

XX 17-OCT-2002.

XX 29-MAR-2002; 2002WO-US009776.

XX 04-APR-2001; 2001WO-US010981.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nelson LM, Tong Z;

XX WPI; 2003-058494/05.

DR P-PSDB; AAE31747.

XX New isolated variant MATER proteins and nucleic acids, useful for
PT diagnosing, prognosing and treating infertility and reduced fertility,
PT and as contraceptive agents.

PS Example 2; Page 81-82; 110pp; English.

XX The present invention relates to novel MATER proteins and polynucleotides
CC encoding such proteins. The MATER proteins are essential to fertility.
CC Sequences of the invention are useful for diagnosing, prognosing and
CC treating infertility, reduced fertility and as contraceptive agents. They
CC are also useful in gene therapy. The method is useful for detecting a
CC predisposition to or pre-symptomatic screening of an individual for
CC infertility or reduced fertility. The present sequence is human MATER
CC cDNA fragment 2

XX Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.46e-184 Length: 1075
Score: 196.00 Matches: 252

Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 1
Query Match: 14.58% Indels: 2
DB: 7 Gaps: 0

US-10-066-521-6 (1-1344) x AAD49000 (1-1075)

Qy 851 LeuAlaSerAlaLeuValSerAsnArgSerLeuThrHisLeuCysLeuSerAsnAsnSer 870
Db 2 CTGGCCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGCTATCAACAACAGC 61
Qy 871 LeuGlyAsnGluGlyValAsnLeuLeuCysArgSerMetArgLeuProHisCysSerLeu 890
Db 62 CTGGGGAACGAAGGTGTAATCTACTGTGTGTCATCCATGAGGCTTCCCCACTGTAGTCTG 121
Qy 891 GlnArgLeuMetLeuAsnGlnCysHisLeuAspThrAlaGlyCysGlyPhe-LeuAlaLe 910
Db 122 CAGAGGCTGATGCTGAATCAGTGCACCTCGACAGCGCTGGCTGTGG-TTCTCTTGCAC 180
Qy 910 uAlaLeuMetGlyAsnSerTrpLeuThrHisLeuSerLeuSerMetAsnProValGluAs 930
Db 181 TGGCGCTTATGGGTAACCTCATGGCTGACGCACCTGAGCCTTAGCATGAACCTGTGAAGA 240
Qy 930 pAsnGlyValLysLeuLeuCysGluValMetArgGluProSerCysHisLeuGlnAspLe 950
Db 241 CAATGGCGTGAAGCTTCTGTGGAGGTCATGAGAGAACCACTCTGTCTCCTCAGGACCT 300
Qy 950 uGluLeuValLysCysHisLeuThrAlaLaCysCysGluSerLeuSerCysValIleSe 970
Db 301 GGAGTTGGTAAATGTCATCTCACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTC 360
Qy 970 ArgSerArgHisLeuLysSerLeuAspLeuThrAspAsnAlaLeuGlyAspGlyGlyVa 990
Db 361 GAGGAGCAGACACCTGAAGAGCCTGGATCTCACGGAACAATGCCCTGGGTGACGGTGGGG 420
Qy 990 lAlaAlaLeuCysGluGlyLeuLysGlnLysAsnSerValLeuThrArgLeuGlyLeuLy 1010
Db 421 TGTGCGCTGTGCGGGGACTGAAGCAAAAGACAGTGTCTGACGAGACTCGGGTTGAA 480
Qy 1010 sAlaCysGlyLeuThrSerAspCysCysGluAlaLeuSerLeuAlaLeuSerCysAsnAr 1030
Db 481 GGCATGTGGACTGACTTCTGATTTGCTGTGAGGCACCTCTCTTGGCCCTTTCTTGCACCG 540
Qy 1030 gHisLeuThrSerLeuAsnLeuValGlnAsnAsnPheSerProLysGlyMetMetLysLe 1050
Db 541 GCATCTGACCAGTCTAAACCTGTGCGAGATAAATCTTCAGTCCCAAGGAATGATGAAGCT 600
Qy 1050 uCysSerAlaPheAlaCysProThrSerAsnLeuGlnIleIleGlyLeuTrpLysTpGl 1070
Db 601 GTGTTGGCCCTTTGCCCTGCCACGTCCTAACTTACAGATTAATTGGGCTGTGGAATGGCA 660
Qy 1070 nTyrProValGlnIleArgLysLeuLeuGluValGlnLeuLeuLysProArgValVa 1090
Db 661 GTACCTGTGCAAAATAAGGAAGCTGTCTGGAGGAAGTGCAGCTACTCAAGCCCCGAGTCGT 720
Qy 1090 lIleAspGlySerTrpHisSerPheAspGluAspArg 1103
Db 721 AATTGACGGTAGTTGGCAATCTTTTGATGAAGATGACCG 760

RESULT 15

ADE36416

ID ADE36416 standard; cDNA; 3108 BP.

XX ADE36416;

XX AC

XX 29-JAN-2004 (first entry)

XX Human PAAD and nucleotide binding protein PAN6 cDNA.

XX cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotrophic;
KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFKappaB activation inhibitor; PAAD domain containing polypeptide;

